
W P E R E H (TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPSrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Feb 26 10:17:57 1998; MasPar time 5.85 Seconds
Tabular output not generated. 540.098 Million cell updates/sec

Title: >US-08-874-460-2
Description: (1-149) from 6808874460.pep
Perfect Score: 1099
Sequence: 1 MNWLLACLIVAGFLGAWAPA.....NPISSSKRVNVLISANSGL 149

Scoring table: PAM 150
Gap 11

Searched: 59021 seqs, 21210388 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot34

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11

Statistics: Mean 42.837; Variance 68.341; scale 0.627

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
. and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description | Pred. No. |
|------------|-------|-------------|--------|-------------|-----------------------|-----------|
| 1 | 119 | 10.8 | 97 | EOTA_HUMAN | EOTAXIN PRECURSOR (EO | 3.75e-06 |
| 2 | 118 | 10.7 | 122 | MILG_MOUSE | MACROPHAGE INFLAMMATO | 5.67e-06 |
| 3 | 116 | 10.6 | 116 | C10_MOUSE | C10 PROTEIN PRECURSOR | 1.29e-05 |
| 4 | 112 | 10.2 | 96 | EOTA_CAVPO | EOTAXIN PRECURSOR (EO | 6.50e-05 |
| 5 | 107 | 9.7 | 92 | MILIA_RAT | MACROPHAGE INFLAMMATO | 4.70e-04 |
| 6 | 105 | 9.6 | 92 | MILIA_MOUSE | MACROPHAGE INFLAMMATO | 1.02e-03 |
| 7 | 105 | 9.6 | 99 | MCP2_BOVIN | MONOCYTE CHEMOTACTIC | 1.02e-03 |
| 8 | 102 | 9.3 | 92 | MILIA_HUMAN | MACROPHAGE INFLAMMATO | 3.23e-03 |
| 9 | 102 | 9.3 | 99 | MCP2_PIG | MONOCYTE CHEMOTACTIC | 3.23e-03 |
| 10 | 101 | 9.2 | 148 | MCP1_MOUSE | MONOCYTE CHEMOTACTIC | 4.71e-03 |
| 11 | 100 | 9.1 | 93 | MILIO_HUMAN | TONSILLAR LYMPHOCYTE | 6.86e-03 |
| 12 | 98 | 8.9 | 92 | MILIA_MOUSE | MACROPHAGE INFLAMMATO | 1.44e-02 |
| 13 | 98 | 8.9 | 618 | WEEL_DROME | WEEL-LIKE PROTEIN KIN | 1.44e-02 |
| 14 | 97 | 8.8 | 246 | RL7_DICDI | 60S RIBOSOMAL PROTEIN | 2.09e-02 |
| 15 | 97 | 8.8 | 476 | BIP_RAT | BONE INDUCING PROTEIN | 2.09e-02 |
| 16 | 96 | 8.7 | 77 | MCP2_HUMAN | MONOCYTE CHEMOTACTIC | 3.01e-02 |
| 17 | 94 | 8.6 | 114 | LTN_RAT | LYMPHOTACTIN PRECURSO | 6.21e-02 |
| 18 | 92 | 8.4 | 91 | SISD_HUMAN | T-CELL SPECIFIC RANGE | 1.27e-01 |
| 19 | 91 | 8.3 | 99 | MCP2_HUMAN | MONOCYTE CHEMOTACTIC | 1.80e-01 |
| 20 | 91 | 8.3 | 847 | MDOH_ECOLI | PERIPLASMIC GLUCANS B | 1.80e-01 |
| 21 | 90 | 8.2 | 114 | LTN_MOUSE | LYMPHOTACTIN PRECURSO | 2.56e-01 |
| 22 | 89 | 8.1 | 151 | YAM5_SCHPO | HYPOTHETICAL 16.7 KD | 3.62e-01 |

ND

| | | | | | | | |
|----|----|-----|------|----|------------|-----------------------|----------|
| 23 | 88 | 8.0 | 288 | 2 | CD80_HUMAN | T LYMPHOCYTE ACTIVATI | 5.10e-01 |
| 24 | 88 | 8.0 | 1631 | 10 | VIT6_CAEEL | VITELLOGENIN 6 PRECUR | 5.10e-01 |
| 25 | 87 | 7.9 | 91 | 9 | SISD_MOUSE | T-CELL SPECIFIC RANGE | 7.17e-01 |
| 26 | 87 | 7.9 | 92 | 6 | MILB_RAT | MACROPHAGE INFLAMMATO | 7.17e-01 |
| 27 | 87 | 7.9 | 92 | 9 | SISD_RAT | T-CELL SPECIFIC RANGE | 7.17e-01 |
| 28 | 87 | 7.9 | 448 | 10 | VDR_COTJA | VITAMIN D3 RECEPTOR (| 7.17e-01 |
| 29 | 87 | 7.9 | 731 | 5 | HRPM_PSESY | HRPM PROTEIN. | 7.17e-01 |
| 30 | 86 | 7.8 | 125 | 6 | MCPI_RABIT | MONOCYTE CHEMOTACTIC | 1.00e+00 |
| 31 | 86 | 7.8 | 359 | 11 | YKH5_CAEEL | HYPOTHETICAL 41.1 KD | 1.00e+00 |
| 32 | 86 | 7.8 | 617 | 9 | THRE_RAT | PROTHROMBIN PRECURSOR | 1.00e+00 |
| 33 | 86 | 7.8 | 618 | 9 | THRB_MOUSE | PROTHROMBIN PRECURSOR | 1.00e+00 |
| 34 | 85 | 7.7 | 92 | 6 | MILB_HUMAN | MACROPHAGE INFLAMMATO | 1.40e+00 |
| 35 | 85 | 7.7 | 97 | 3 | EOTA_MOUSE | EOTAXIN PRECURSOR (EO | 1.40e+00 |
| 36 | 85 | 7.7 | 99 | 6 | MCPI_HUMAN | MONOCYTE CHEMOTACTIC | 1.40e+00 |
| 37 | 85 | 7.7 | 614 | 1 | AH11_ARAHY | ALLERGEN ARA H I, CLO | 1.40e+00 |
| 38 | 83 | 7.6 | 105 | 11 | YG44_YEAST | HYPOTHETICAL 11.3 KD | 2.71e+00 |
| 39 | 83 | 7.6 | 280 | 7 | ODO1_COXBU | 2-OXOGLUTARATE DEHYDR | 2.71e+00 |
| 40 | 83 | 7.6 | 315 | 10 | VN35_ROT11 | NONSTRUCTURAL RNA-BIN | 2.71e+00 |
| 41 | 83 | 7.6 | 1084 | 6 | MYSS_RABIT | MYOSIN HEAVY CHAIN, S | 2.71e+00 |
| 42 | 84 | 7.6 | 1571 | 1 | ATC5_YEAST | PROBABLE CALCIUM-TRAN | 1.95e+00 |
| 43 | 83 | 7.6 | 3110 | 4 | HD_RAT | HUNTINGTIN (HUNTINGTO | 2.71e+00 |
| 44 | 83 | 7.6 | 3119 | 4 | HD_MOUSE | HUNTINGTIN (HUNTINGTO | 2.71e+00 |
| 45 | 82 | 7.5 | 1102 | 6 | MYSC_CHICK | MYOSIN HEAVY CHAIN, C | 3.75e+00 |

ALIGNMENTS

RESULT 1

ID EOTA_HUMAN STANDARD; PRT; 97 AA.

AC P51671; P50877;

DT 01-OCT-1996 (REL. 34, CREATED)

DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)

DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)

DE EOTAXIN PRECURSOR (EOSINOPHIL CHEMOTACTIC PROTEIN).

GN SCY11

OS HOMO SAPIENS (HUMAN).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;

OC EUTHERIA; PRIMATES.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE; 96181758.

RA GARCIA-ZEPEDA E.A., ROTHENBERG M.E., OWNBEY T.R., LEDER P.,

RA LUSTER A.D.;

RL NAT. MED. 2:449-456(1996).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE; 96189937.

RA PONTATH P.D., QIN S., RINGLER D.J., CLARK-LEWIS I., WANG J., KASSAM N.,

RA SMITH H., SHI X., GONZALO J.A., NEWMAN W., GUTIERREZ-RAMOS J.C.,

RA MACKAY C.R.;

RL J. CLIN. INVEST. 97:604-612(1996).

RN [3]

RP SEQUENCE OF 1-59 FROM N.A.

RC TISSUE-FORESKIN;

RX MEDLINE; 96374440.

RA BARTELS J., SCHLUETER C., RICHTER E., NOSO N., KULKE R.,

RA CHRISTOPHERS E., SCHROEDER J.M.;

RL BIOCHEM. BIOPHYS. RES. COMMUN. 225:1045-1051(1996).

CC -!- FUNCTION: IN RESPONSE TO THE PRESENCE OF ALLERGENS, THIS PROTEIN

CC DIRECTLY PROMOTES THE ACCUMULATION OF EOSINOPHILS, A PROMINENT

CC FEATURE OF ALLERGIC INFLAMMATORY REACTIONS.

CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.

CC -!- PTM: O-GLYCOSYLATED (PROBABLE).

CC -!- SIMILARITY: BELONGS TO THE INTERCINE BETA FAMILY (SMALL CYTOKINE

CC C-C) (CHEMOKINE CC).

DR EMBL; U46573; G1280141; -

DR EMBL; U34780; G1185440; -

DR EMBL; 269291; E221070; -

DR MIM; 601156; -

DR EOTAXIN; CHEMOTAXIS; GLYCOPROTEIN; SIGNAL.

FT SIGNAL 1 23 POTENTIAL.

FT CHAIN 24 97 EOTAXIN PRECURSOR.

FT DISULFID 32 57 BY SIMILARITY.

| | |
|--|---|
| CC | C-C) (CHEMOKINE CC). |
| DR | EMBL; U22414; G790633; -. |
| DR | EMBL; U06435; G459150; -. |
| KW | CTKINE; CHEMOTAXIS; INFLAMMATORY RESPONSE; SIGNAL. |
| FT | SIGNAL 1 23 |
| FT | CHAIN 24 92 MACROPHAGE INFLAMMATORY PROTEIN 1-ALPHA. |
| FT | DLSULFID 34 57 BY SIMILARITY. |
| FT | DLSULFID 35 73 BY SIMILARITY. |
| FT | CONFLICT 6 6 A -> T (IN REF. 2). |
| FT | CONFLICT 57 57 C -> W (IN REF. 2 AND 3). |
| SQ | SEQUENCE 92 AA; 10335 MW; F48CF89F CRC32; |
| Query Match 9.7%; Score 107; DB 6; Length 92; Best Local Similarity 41.0%; Pred. No. 4.70e-04; Matches 16; Conservative 9; Mismatches 13; Indels 1; Gaps | |
| Dd | 52 etslscsqpvif-ltknrqicadpketwvqvitele 89 : : : : : : : |
| Qy | 53 EVSGCNLPAIFLYPKRHRKVCNGPNKSREVQRAMKLDD 91 : : : : : : : |
| RESULT | 6 STANDARD; PRT; 92 AA. |
| AC | P14097; |
| DT | 01-JAN-1990 (REL. 13, CREATED) |
| DT | 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE) |
| DT | 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE) |
| DE | MACROPHAGE INFLAMMATORY PROTEIN 1-BETA PRECURSOR (MIP-1-BETA) (H400 |
| DE | PROTEIN) (SIS-GNMM) (AC12). |
| GN | SC1A4 OR MIPB. |
| OS | MUS MUSCULUS (HOUSE). |
| OC | EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA; |
| OC | EUTHERIA; RODENTIA. |
| [1] | |
| RP | SEQUENCE FROM N.A. |
| RX | MEDLINE; 89067830. |
| RA | SHERY B., TEKAMP-OLSON P., GALLEGOS C., BAUER D., DAVATELIS G., |
| RA | WOLPE S.D., MAGIARZ F., COIT D., CERAMI A.; |
| RA | J. EXP. MED. 168:2251-2259(1988). |
| [2] | |
| RP | SEQUENCE FROM N.A. |
| RX | MEDLINE; 89093358. |
| RA | BROWN K.D., ZURAWSKI S.M., MOSMANN T.R., ZURAWSKI G.; |
| RA | J. IMMUNOL. 142:679-687(1989). |
| [3] | |
| RP | SEQUENCE FROM N.A. |
| RC | STRAIN-DBA/J; TISSUE-LIVER; |
| RA | DAUBERSIES P., LEPRETRE F., BAILLEUL B., GROVE M., PRAGNELL I., |
| RA | PLUMB M.A.; |
| RL | SUBMITTED (OCT-1991) TO EMBL/GENBANK/DBJ DATA BANKS. |
| CC | -1- FUNCTION: MONOKINE WITH INFLAMMATORY AND CHEMOKINETIC PROPERTIES. |
| CC | -1- SUBUNIT: HOMODIMER (BY SIMILARITY). |
| CC | -1- SIMILARITY: BELONGS TO THE INTERCRINE BETA FAMILY (SMALL CYTOKINE |
| CC | C-C) (CHEMOKINE CC). |
| DR | EMBL; M23503; G533245; -. |
| DR | EMBL; M35590; G199697; -. |
| DR | EMBL; X62502; G53127; -. |
| DR | PIR; C30552; C30552. |
| DR | PIR; JLO088; JLO088. |
| DR | HSSP; P13236; IHUM. |
| DR | PROSITE; PS00472; SMALL_CYTOKINES_CC. |
| KW | CTKINE; CHEMOTAXIS; INFLAMMATORY RESPONSE; SIGNAL. |
| FT | SIGNAL 1 23 |
| FT | CHAIN 24 92 MACROPHAGE INFLAMMATORY PROTEIN 1-BETA. |
| FT | DLSULFID 34 58 BY SIMILARITY. |
| FT | DLSULFID 35 74 BY SIMILARITY. |
| FT | CONFLICT 75 75 A -> P (IN REF. 1). |
| FT | CONFLICT 79 79 E -> Q (IN REF. 1). |
| FT | CONFLICT 88 88 D -> H (IN REF. 1). |
| SQ | SEQUENCE 92 AA; 10168 MW; C543B91F CRC32; |
| Query Match 9.6%; Score 105; DB 6; Length 92; Best Local Similarity 32.5%; Pred. No. 1.02e-03; | |

[illegible]

Q

WATERMAN (TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.

Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.

Distribution rights by IntelliGenetics, Inc.

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Thu Feb 26 10:48:37 1998; MasPar time 550.37 Seconds

Tabular output not generated. 887.394 Million cell updates/sec

Title: >US-08-874-460-1

Description: (1-989) from US08874460.seq

Perfect Score: 989

N.A. Sequence: 1 CGCGGGCGCATCGCTCCCT.....AAAAA.....989

Comp: GGCGCGCGTAGTCGAGGGA.....TTTTTTTTTTTTTTTT

Scoring table: TABLE default

Gap 6

Mmatch STD : Dbase 0; Query 0

Searched: 665703 seqs, 246912890 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: EST-A

1:EST1 2:EST2 3:EST3 4:EST4 5:EST5 6:EST6 7:EST7 8:EST8
9:EST9 10:EST10 11:EST11 12:EST12 13:EST13 14:EST14
15:EST15 16:EST16 17:EST17 18:EST18 19:EST19 20:EST20
21:EST21 22:EST22 23:EST23 24:EST24 25:EST25 26:EST26
27:EST27 28:EST28 29:EST29 30:EST30 31:EST31 32:EST32
33:EST33 34:EST34 35:EST35 36:EST36 37:EST37 38:EST38
39:EST39 40:EST40 41:EST41 42:EST42 43:EST43 44:EST44
45:EST45 46:EST46 47:EST47 48:EST48 49:EST49 50:EST50
51:EST51 52:EST52 53:EST53 54:EST54 55:EST55 56:EST56
57:EST57 58:EST58 59:EST59 60:EST60 61:EST61 62:EST62
63:EST63 64:EST64 65:EST65 66:EST66 67:EST67 68:EST68
69:EST69 70:EST70 71:EST71 72:EST72 73:EST73 74:EST74
75:EST75 76:EST76 77:EST77 78:EST78 79:EST79 80:EST80
81:EST81 82:EST82 83:EST83 84:EST84 85:EST85 86:EST86
87:EST87 88:EST88 89:EST89 90:EST90 91:EST91 92:EST92
93:EST93 94:EST94 95:EST95 96:EST96 97:EST97 98:EST98
EST-B

Database: EST-B

99:EST99 100:EST100 101:EST101 102:EST102 103:EST103
104:EST104 105:EST105 106:EST106 107:EST107 108:EST108
109:EST109 110:EST110 111:EST111 112:EST112 113:EST113
114:EST114 115:EST115 116:EST116 117:EST117 118:EST118
119:EST119 120:EST120 121:EST121 122:EST122 123:EST123
124:EST124 125:EST125 126:EST126 127:EST127 128:EST128
129:EST129 130:EST130 131:EST131 132:EST132 133:EST133
134:EST134 135:EST135 136:EST136 137:EST137 138:EST138
139:EST139 140:EST140 141:EST141 142:EST142 143:EST143
144:EST144 145:EST145 146:EST146 147:EST147 148:EST148
149:EST149 150:EST150 151:EST151 152:EST152 153:EST153
154:EST154 155:EST155 156:EST156 157:EST157 158:EST158
159:EST159 160:EST160 161:EST161 162:EST162 163:EST163
164:EST164 165:EST165 166:EST166 167:EST167 168:EST168
169:EST169 170:EST170 171:EST171 172:EST172 173:EST173
174:EST174 175:EST175 176:EST176 177:EST177 178:EST178

179:EST179 180:EST180 181:EST181 182:EST182 183:EST183
184:EST184 185:EST185 186:EST186 187:EST187 188:EST188
189:EST189 190:EST190 191:EST191 192:EST192 193:EST193
194:EST194 195:EST195 196:EST196

Statistics: Mean 11.471; Variance 3.780; scale 3.035

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description | Pred. No. |
|------------|-------|-------------|---------|----------|-----------------------|-----------|
| C 1 | 366 | 37.0 | 407 96 | N73958 | za74f04.s1 Homo sapie | 0.00e+00 |
| C 2 | 160 | 16.2 | 379 185 | SS284185 | S.scrofa MRNA; expres | 8.43e-153 |
| C 3 | 99 | 10.0 | 378 135 | W91616 | MFA.F02.091.A MFA adu | 1.96e-79 |
| C 4 | 38 | 3.8 | 172 189 | AA046474 | zf48g09.s1 Soares ret | 5.32e-13 |
| C 5 | 36 | 3.6 | 275 144 | AA035211 | ze24a03.s1 Soares fet | 3.78e-11 |
| C 6 | 36 | 3.6 | 333 91 | H75269 | yu07f06.r1 Homo sapie | 3.78e-11 |
| C 7 | 36 | 3.6 | 392 150 | T18721 | 5C04H03-T7 membrane-f | 3.78e-11 |
| C 8 | 36 | 3.6 | 450 135 | W95880 | ze08c04.s1 Homo sapie | 3.78e-11 |
| C 9 | 35 | 3.5 | 146 76 | N26762 | yz96d02.s1 Homo sapie | 3.07e-10 |
| C 10 | 35 | 3.5 | 250 87 | N67550 | yz35f05.s1 Homo sapie | 3.07e-10 |
| C 11 | 35 | 3.5 | 263 175 | AA180346 | zpl5c07.s1 Stratagene | 3.07e-10 |
| C 12 | 35 | 3.5 | 304 22 | R25629 | yh45g06.r1 Homo sapie | 3.07e-10 |
| C 13 | 35 | 3.5 | 329 177 | AA189446 | mt87h06.r1 Soares mou | 3.07e-10 |
| C 14 | 35 | 3.5 | 360 178 | AA196465 | zp98h12.s1 Stratagene | 3.07e-10 |
| C 15 | 35 | 3.5 | 438 192 | AA136868 | z101c03.s1 Soares pre | 3.07e-10 |
| C 16 | 35 | 3.5 | 449 132 | W65270 | me06g06.r1 Soares mou | 3.07e-10 |
| C 17 | 35 | 3.5 | 488 86 | N64084 | yz41g12.s1 Homo sapie | 3.07e-10 |
| C 18 | 34 | 3.4 | 193 177 | AA189534 | mt90c08.r1 Soares mou | 2.43e-09 |
| C 19 | 34 | 3.4 | 194 182 | AA210586 | mu71b03.r1 Soares mou | 2.43e-09 |
| C 20 | 34 | 3.4 | 212 24 | R36217 | yh91a10.s1 Homo sapie | 2.43e-09 |
| C 21 | 34 | 3.4 | 222 169 | AA017921 | mb46b08.r1 Soares mou | 2.43e-09 |
| C 22 | 34 | 3.4 | 250 127 | W20545 | mb94a10.r1 Soares mou | 2.43e-09 |
| C 23 | 34 | 3.4 | 257 11 | T95813 | ye45f10.r1 Homo sapie | 2.43e-09 |
| C 24 | 34 | 3.4 | 383 38 | H08710 | yi93h12.s1 Homo sapie | 2.43e-09 |
| C 25 | 34 | 3.4 | 310 40 | R10582 | yf35b11.r1 Homo sapie | 2.43e-09 |
| C 26 | 34 | 3.4 | 343 155 | AA057833 | z195c06.s1 Stratagene | 2.43e-09 |
| C 27 | 34 | 3.4 | 345 25 | R38991 | yd07h01.s1 Homo sapie | 2.43e-09 |
| C 28 | 34 | 3.4 | 352 125 | W14029 | ma64a05.r1 Soares mou | 2.43e-09 |
| C 29 | 34 | 3.4 | 354 100 | N98932 | zb87d03.s1 Homo sapie | 2.43e-09 |
| C 30 | 34 | 3.4 | 362 162 | AA060885 | mj86h03.r1 Soares mou | 2.43e-09 |
| C 31 | 34 | 3.4 | 382 174 | AA173407 | zp02b01.s1 Stratagene | 2.43e-09 |
| C 32 | 34 | 3.4 | 446 191 | AA081389 | zn34e10.s1 Stratagene | 2.43e-09 |
| C 33 | 34 | 3.4 | 456 45 | H18464 | ym43f08.s1 Homo sapie | 2.43e-09 |
| C 34 | 34 | 3.4 | 460 135 | W91591 | MFA.D04.042.A MFA adu | 2.43e-09 |
| C 35 | 34 | 3.4 | 461 178 | AA192637 | zq01h07.s1 Stratagene | 2.43e-09 |
| C 36 | 34 | 3.4 | 465 178 | AA196593 | zq07g05.s1 Stratagene | 2.43e-09 |
| C 37 | 34 | 3.4 | 525 160 | W88002 | mf65c03.r1 Soares mou | 2.43e-09 |
| C 38 | 34 | 3.4 | 571 118 | AA001414 | ze45d08.s1 Soares ret | 2.43e-09 |
| C 39 | 34 | 3.4 | 670 139 | AA032833 | ml38g08.r1 Soares mou | 2.43e-09 |
| C 40 | 33 | 3.3 | 330 135 | W84086 | TJ987 MWAT4 bloodstre | 1.87e-08 |
| C 41 | 33 | 3.3 | 361 165 | AA086043 | z184e01.s1 Stratagene | 1.87e-08 |
| C 42 | 33 | 3.3 | 419 132 | W64439 | md65e03.r1 Soares mou | 1.87e-08 |
| C 43 | 33 | 3.3 | 439 182 | AA206838 | zg80f03.s1 Stratagene | 1.87e-08 |
| C 44 | 33 | 3.3 | 441 133 | W70638 | me34c10.r1 Soares mou | 1.87e-08 |
| C 45 | 33 | 3.3 | 506 153 | W44573 | zc29d03.s1 Soares sen | 1.87e-08 |

ALIGNMENTS

1 N73958 407 bp mRNA EST
za74f04.s1 Homo sapiens CDNA clone 298303 3'
N73958
g1231243
EST.
human clone-298303 primer-ml3 -40 forward library-Soares fetal lung
NbH19W vector-pt7T3D (Pharmacia) with a modified polylinker

19-MAR-1996

Gainesville, FL 32611-0695
ph: 904-392-1928, ext. 301
fax: 904-392-4072
E-mail: robferl@nervm.nerdc.ufl.edu

Seq primer: T7.
Location/Qualifiers
1..392

```

/organism="Zea mays"
/strain="W64A2"
/notes="Vector: ZipLox; Site1: SalI; Site2: NotI; ds-cDNA
was prepared from oligo-dT selected mRNA by priming with a
NotI oligo- dT oligomer and then adding the second strand
to RNase-nicked DNA:RNA hybrid with DNA polI. SalI
adaptors were added to the ends, the ds-cDNAs were then
digested with NotI and size-selected. These were
directionally-cloned into the ZipLox phage vector, excised
as plasmids, and then individually analyzed."
/clone="5C04H03"
/clone_lib="membrane-free polysomes from endosperm"
/lab_host="DH10B"
<!-->392
mRNA

```

| | | | | | |
|-----------------------|-------|--------------|------|------------|----------|
| BASE COUNT | 106 a | 97 c | 92 g | 89 t | 8 others |
| ORIGIN | | | | | |
| Query Match | 3.6% | | | | |
| Best Local Similarity | 95.0% | | | | |
| Matches | 38; | Conservative | 0; | Mismatches | 2; |
| Indels | 0; | Indels | 0; | Caps | 0; |

Db 131 gtgtggagtaaaaaaaaaaaaaaaaaaaaaaa 170
||| ||| ||| ||| ||| ||| ||| ||| |||
QY 950 GTGTGGGTCACAAAAAANAANAANAANA 989

| RESULT | 8 | W95880 | 450 bp | EST | 16-JUL-1996 |
|------------|---|------------|--------------------|---------|-------------------------|
| LOCUS | | ze08c04.s1 | Soares fetal heart | NDHH19W | Homo sapiens cDNA clone |
| DEFINITION | | 358374 3' | | | |

| | | |
|-----------|--|--|
| W33680 | Homo sapiens | Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Hominoidea; Homo. |
| g1425787 | 1 (bases 1 to 450) | |
| NID | Willier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R. | |
| KEYWORDS | | |
| EST. | | |
| SOURCE | human. | |
| ORGANISM | | |
| REFERENCE | | |
| AUTHORS | | |
| TITLE | The WashU-Merck EST Project | |
| JOURNAL | Unpublished (1995) | |

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: mob.REGA+ET
High quality sequence stop: 340.

| FEATURES | Source |
|-----------------------|--|
| high quality sequence | Genbank: U00166 |
| Location/Qualifiers | 1..450 |
| | /organism="Homo sapiens" |
| | /note="Organ: heart; Vector: pT73D (Pharmacia) with a modified polylinker; Site:1: Not I; Site:2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer TGTTCACCATCGAAGTCGGGCGGCCGATCTTTTTTTTTTTT 3'] |

```

double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by
M.Fatima Bonaldo. This library was constructed from the
same fetus as the fetal lung library, Soares fetal lung
NbHL19W."
/clone="358374"
/clone_lib="Soares fetal heart NbHH19W"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DHL0B (ampicillin resistant)"
complement(<i...>450)
91 a 104 c 121 g 129 t 5 others
mRNA
BASE COUNT
ORIGIN

```

```
Query Match      3.6%; Score 36; DB 135; Length 450;
Best Local Similarity 76.5%; Pred. No. 3.78e-11;
Matches 52; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
```

[illegible]

Db 67 gctgtgga 74
Cp 929 GTTATAGA 922

| | |
|------------|--|
| RESULT | 9 |
| LOCUS | N26762 146 bp mRNA EST |
| DEFINITION | YX96D02.s1 Homo sapiens cDNA clone 269571 3'. |
| ACCESSION | N26762 |
| NID | g1141110 |
| KEYWORDS | EST. |
| SOURCE | Homo clone-269571 primer-m13 -40 forward library=Soares.mala |

human clone-269571 primer=ml3 -40 forward library=Soares melanocyte 2NBHM vector=pf7T3D (Pharmacia) with a modified polylinker host=DH10B (ampicillin resistant) Rsite1=Not I Rsite2=Eco RI Male. 1st strand cDNA was primed with a Not I - oligo(dT) primer 15'-TGTTACCAATCTGAAGTGGAGCGCCGCGCATTTTTTTTTTTTTTTT-3', double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified p7T3 vector (Pharmacia). Library constructed by Bento Soares and M.Fatima Bonaldo. RNA from normal foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P. Albino.

ORGANISM

Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 146)

REFERENCE

AUTHORS

Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rife, L., Rohlfing, T., Soares, M., Tan, F., Trevisakis, E., Waterston, R., Williamson, A., Wohlmann, P., and Wilson, R.

TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)

Contact: Wilson RK

Contact: WILSON RN
 WashU-Merck EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810

Email: estefawatson.wustl.edu
 High quality sequence stops: 81
 Source: IMAGE Consortium, LLNL
 This clone is available royalty
 IMAGE Consortium (info@image.ll
 Location/Qualifiers

FEATURES

WQREH (TW)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1995, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPSrch_on n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Thu Feb 26 10:59:09 1998; MasPar time 301.05 Seconds
Tabular output not generated. 926.471 Million cell updates/sec

Title: >US-08-874-460-1
Description: (1-989) from US08874460.seq
Perfect Score: 989
N.A. Sequence: 1 CGCGCGGCATCAGTCCTCCT.....AAAAA.....AAAAA 989
Comp: GCGCGCCGTCAGTCGAGGGA.....TTTTTTTTTTTTTTTTTTTT

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 397346 seqs, 141010104 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: EST-C
1:EST197 2:EST198 3:EST199 4:EST200 5:EST201 6:EST202
7:EST203 8:EST204 9:EST205 10:EST206 11:EST207 12:EST208
13:EST209 14:EST210 15:EST211 16:EST212 17:EST213
18:EST214 19:EST215 20:EST216 21:EST217 22:EST218
23:EST219 24:EST220 25:EST221 26:EST222 27:EST223
28:EST224 29:EST225 30:EST226 31:EST227 32:EST228
33:EST229 34:EST230 35:EST231 36:EST232 37:EST233
38:EST234 39:EST235 40:EST236 41:EST237 42:EST238
43:EST239 44:EST240 45:EST241 46:EST242 47:EST243
48:EST244 49:EST245 50:EST246 51:EST247 52:EST248
53:EST249 54:EST250 55:EST251 56:EST252 57:EST253
58:EST254 59:EST255 60:EST256 61:EST257 62:EST258
63:EST259 64:EST260 65:EST261 66:EST262 67:EST263
68:EST264 69:EST265 70:EST266 71:EST267 72:EST268
73:EST269 74:EST270 75:EST271 76:EST272 77:EST273
78:EST274 79:EST275 80:EST276 81:EST277 82:EST278
83:EST279 84:EST280 85:EST281 86:EST282 87:EST283
88:EST284 89:EST285 90:EST286 91:EST287 92:EST288
93:EST289 94:EST290 95:EST291 96:EST292 97:EST293
98:EST294
EST-D
99:EST295 100:EST296 101:EST297 102:EST298 103:EST299
104:EST300 105:EST301 106:EST302 107:EST303 108:EST304
109:EST305 110:EST306 111:EST307 112:EST308 113:EST309
114:EST310 115:EST311 116:EST312 117:EST313 118:EST314
119:EST315

Statistics: Mean 11.352; Variance 3.381; scale 3.358
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description | Pred. No. |
|------------|-------|-------------|--------|-------|-------------|-----------------------|
| 1 | 286 | 28.9 | 291 | 25 | AA295814 | EST101116 Thymus III |
| 2 | 261 | 26.4 | 279 | 25 | AA295958 | EST101182 Thymus III |
| 3 | 254 | 25.7 | 274 | 25 | AA295925 | EST101148 Thymus III |
| 4 | 230 | 23.3 | 256 | 25 | AA295945 | EST101169 Thymus III |
| 5 | 118 | 11.9 | 444 | 49 | AA277792 | vb68a11.r1 Barstead m |
| 6 | 37 | 3.7 | 231 | 47 | AA277068 | va69g08.r1 Soares m |
| 7 | 37 | 3.7 | 428 | 92 | C14427 | Human fetal brain CDN |
| 8 | 37 | 3.7 | 432 | 11 | AA237207 | mx18e02.r1 Soares mou |
| 9 | 36 | 3.6 | 224 | 93 | C16161 | Human aorta CDNA 5'-e |
| 10 | 36 | 3.6 | 325 | 1 | AA276034 | mt04e07.r1 Soares mou |
| 11 | 36 | 3.6 | 331 | 44 | AA260469 | va44d04.r1 Soares mou |
| 12 | 36 | 3.6 | 430 | 92 | C14394 | Human fetal brain CDN |
| 13 | 35 | 3.5 | 403 | 47 | AA270895 | vb72e11.r1 Soares mou |
| 14 | 35 | 3.5 | 413 | 46 | AA267728 | mx66e03.r1 Soares mou |
| 15 | 34 | 3.4 | 94 | 65 | AA396012 | vb42e03.r1 Soares mou |
| 16 | 34 | 3.4 | 97 | 12 | AA241100 | mv19h03.r1 Barstead m |
| 17 | 34 | 3.4 | 128 | 49 | AA276953 | vc46g11.r1 Soares mou |
| 18 | 34 | 3.4 | 179 | 46 | AA267708 | mx68h08.r1 Soares mou |
| 19 | 34 | 3.4 | 197 | 6 | AA221995 | mw16e06.r1 Soares mou |
| 20 | 34 | 3.4 | 199 | 6 | AA219890 | mw64e06.r1 Soares mou |
| 21 | 34 | 3.4 | 211 | 48 | AA273932 | vb99d07.r1 Soares mou |
| 22 | 34 | 3.4 | 216 | 43 | AA259890 | va53a02.r1 Soares mou |
| 23 | 34 | 3.4 | 227 | 6 | AA220092 | mv63a11.r1 Soares mou |
| 24 | 34 | 3.4 | 228 | 89 | AA36703 | zv59e06.s1 Soares tes |
| 25 | 34 | 3.4 | 246 | 46 | AA368648 | va43c04.r1 Soares mou |
| 26 | 34 | 3.4 | 259 | 8 | AA230395 | mw04e12.r1 Soares mou |
| 27 | 34 | 3.4 | 269 | 92 | C14215 | Human fetal brain CDN |
| 28 | 34 | 3.4 | 295 | 46 | AA267875 | va18g08.r1 Soares mou |
| 29 | 34 | 3.4 | 349 | 113 | AA445723 | vc62e06.s1 Knowles So |
| 30 | 34 | 3.4 | 366 | 11 | AA238803 | mx92e02.r1 Soares mou |
| 31 | 34 | 3.4 | 403 | 93 | C16061 | Human aorta CDNA 5'-e |
| 32 | 34 | 3.4 | 426 | 80 | AA402661 | zu49f04.s1 Soares ova |
| 33 | 34 | 3.4 | 429 | 13 | AA244918 | mv30e05.r1 Barstead m |
| 34 | 33 | 3.3 | 94 | 12 | AA241234 | mu95e06.r1 Soares mou |
| 35 | 33 | 3.3 | 202 | 15 | W40015 | 455 Mouse VM CDNA lib |
| 36 | 33 | 3.3 | 267 | 49 | AA285567 | vb88f11.r1 Soares mou |
| 37 | 33 | 3.3 | 329 | 110 | RICC1646B | Rice CDNA, partial se |
| 38 | 33 | 3.3 | 355 | 1 | AA176022 | mt04d03.r1 Soares mou |
| 39 | 33 | 3.3 | 371 | 110 | RICR2825A | Rice CDNA, partial se |
| 40 | 33 | 3.3 | 379 | 97 | AA278864 | zs77h11.s1 Soares NbH |
| 41 | 33 | 3.3 | 381 | 7 | AA277123 | zr22d05.s1 Stratagene |
| 42 | 33 | 3.3 | 431 | 94 | AA121171 | zl18g05.s1 Stratagene |
| 43 | 33 | 3.3 | 433 | 101 | AA440484 | LD13087.5Prime LD Dro |
| 44 | 33 | 3.3 | 471 | 47 | AA271476 | va83b12.r1 Soares mou |
| 45 | 33 | 3.3 | 532 | 77 | AA042856 | zk52e08.r1 Soares pre |

ALIGNMENTS

RESULT 1
LOCUS AA295814 291 bp mRNA EST 18-APR-1997
DEFINITION EST101116 Thymus III Homo sapiens CDNA 5' end.
ACCESSION AA295814
NID G1948238
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae;
Homo.
REFERENCE
1 (bases 1 to 291)
Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkley,P.S., Jr.,
Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,

```

ACCESSION   AA295925
NID         91948269
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens

REFERENCE   1 (bases 1 to 274)
AUTHORS    Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
            Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
            White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
            Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
            Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
            Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr.,
            Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
            Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
            Phillips,C.A., Ryder,S.E., Scott,J.L., Saudke,D.M., Shirley,R.,
            Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
            Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
            Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
            He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
            Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H.,
            Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
            Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
            Fraser,C.M. and Venter,J.C.
            Initial assessment of human gene diversity and expression patterns
            based upon 83 million nucleotides of cDNA sequence
            Nature 377 (6547 Suppl), 3-174 (1995)
JOURNAL     96026280
MEDLINE
COMMENT     Other ESTs: THC148619
            Contact: Kerlavage, AR
            Bioinformatics
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850 USA
            Tel: 3018699056
            Fax: 3018699423
            Email: arkerlavet@tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please check the TIGR Human Gene
            Index (http://www.tigr.org/tdb/hgi/hgi.html)
            Seq primer: M13 Reverse.
            Location/Qualifiers
                1..274
                /organism="Homo sapiens"
                /note="Organ: thymus; Vector: pBluescript SK-; Site_1:
                EcoRI; Site_2: XhoI"
                /clone_lib="Thymus III"
                /dev_stage="adult"
                <1..>274

BASE COUNT  60 a 79 c 57 g 71 t 7 others
ORIGIN
Query Match 25.7%; Score 254; DB 25; Length 274;
Best Local Similarity 96.7%; Pred. No. 9,84e-300;
Matches 261; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

Db 1 ccctcatcgtctaaagaagtttaagttctggaactcccaagttatcatcgtcccaagtttag 60
Qy 414 CCCTCATCGTGTAAAGAAGTTTGAAGTCTGGAACCTCCAAGTTATCATCGTCCCAAGTTTAG 473

Db 61 caatccccatcagcagcaagaagaattntccctccctgatcatcagctaattcaggact 120
Qy 474 CAATCCCATCAGCAGCAGCAAGGAATCTCCCTCCCTGATATCAGCTAAATTCAGGACT 533

Db 121 gtaagccggctattntgggtccatcgccagcagagggcgcgatcttttcccgataa 180
Qy 534 GTGACCCGGCTCATTTCTGGGCTCCATCGGCACAGAGAGGG-CCGGATCTCTTCCGATTA 592

Db 181 aaccgtcgccctacagaccagctgtccccacgncctntnttttgggtcaagctcttaat 240
Qy 593 AACCGTCGCCCTACAGACCCACGCTGTCCCCACGCCCTCTCTCTTTGGGTCAAGTCTTAAT 652

```

```

Db 241 ccctgcacctgaagttggtctctccctctgca 270
Qy 653 CCCTGCACCTGAGTTGGTCTCTCCCTCTGCA 682

RESULT      4
LOCUS       AA295945      256 bp      mRNA      EST      18-APR-1997
DEFINITION EST101169 Thymus III Homo sapiens cDNA 5' end.
ACCESSION   AA295945
NID         91948290
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens

REFERENCE   1 (bases 1 to 256)
AUTHORS    Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
            Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
            White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
            Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
            Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
            Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr.,
            Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
            Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
            Phillips,C.A., Ryder,S.E., Scott,J.L., Saudke,D.M., Shirley,R.,
            Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
            Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
            Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
            He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
            Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H.,
            Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
            Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
            Fraser,C.M. and Venter,J.C.
            Initial assessment of human gene diversity and expression patterns
            based upon 83 million nucleotides of cDNA sequence
            Nature 377 (6547 Suppl), 3-174 (1995)
JOURNAL     96026280
MEDLINE
COMMENT     Other ESTs: THC148619
            Contact: Kerlavage, AR
            Bioinformatics
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850 USA
            Tel: 3018699056
            Fax: 3018699423
            Email: arkerlavet@tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please check the TIGR Human Gene
            Index (http://www.tigr.org/tdb/hgi/hgi.html)
            Seq primer: M13 Reverse.
            Location/Qualifiers
                1..256
                /organism="Homo sapiens"
                /note="Organ: thymus; Vector: pBluescript SK-; Site_1:
                EcoRI; Site_2: XhoI"
                /clone_lib="Thymus III"
                /dev_stage="adult"
                <1..>256

BASE COUNT  70 a 66 c 61 g 56 t 3 others
ORIGIN
Query Match 23.3%; Score 230; DB 25; Length 256;
Best Local Similarity 97.3%; Pred. No. 2,47e-266;
Matches 249; Conservative 0; Mismatches 4; Indels 3; Gaps 1;

Db 1 gtgggaacccccaaagcaggaggtgcagagccatgaagctctcgtgctcgaata 60
Qy 311 GTGGGAACCCCAAAAGCAGGAGGTGCAGAGGCCATGAAGCTCTCTGGATGCTCGAATA 370

Db 61 aggttttttcaaaagctccaccacaacacgcagaccttcccaaggccctcatgctgaa 120
Qy 371 AGGTTTTTGCAGAAGCTCCACCACCAACACGACGACCTTCCCAAG---GCCCTCATGCTGTAA 427

```

1 (sites)
Fujiwara, T., Hirano, H., Hishigaki, H., Horie, M., Kawai, A., Kuga, Y.,
Kyushiki, H., Nagata, M., Okuno, S., Ozaki, K., Shimizu, F., Shimada, Y.,
Shinomiya, H., Suzuki, M., Takaichi, A., Takeda, S., Watanabe, T.,
Maekawa, H., Nakamura, Y. and Takahashi, E.
Otsuka cDNA project
Unpublished (1996)
2 (bases 1 to 224)
Fujiwara, T.
Direct Submission
Submitted (13-MAY-1996) to the DDBJ/EMBL/GenBank databases. Tsutomu
Fujiwara, Otsuka GEN Research Institute, Otsuka Pharmaceutical
Co., Ltd; 463-10 kagasuno Kawachi-cho, Tokushima, Tokushima 771-01,
Japan (Tel:+81-886-65-2888, Fax:+81-886-37-1035)

Unpublished (1996)

Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicinep
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:463492
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 353.
Location/Qualifiers
1..403
/organism="Mus musculus"
/strain="C57BL/6J"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with Not I - oligo(dT) primer [5',
TGTTACCAATCGAAGTCGAGCGCGCTATTTTTTTTTTTT 3'],
on total mouse RNA [provided by Minoru Ko, Wayne State
Univ.]; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
Library went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
/clone="762572"
/clone_lib="Soares mouse 3NME12 5"
/sex="unknown"
/dev_stage="12.5dpc total fetus"
/lab_host="DH10B"
<1..>403
132 a 79 c 97 g 95 t

mRNA
BASE COUNT 132 a 79 c 97 g 95 t
ORIGIN

Query Match 3.5%; Score 35; DB 47; Length 403;
Best Local Similarity 94.9%; Pred. No. 3.45e-12;
Matches 37; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

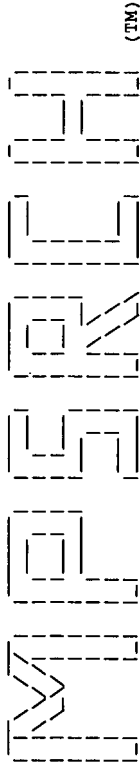
Db 330 tgatggtgtaaaaaaiaaaaaaaaaaaaaaaaaaaaaa 368
||||| | ||||||| ||||||| ||||||| |||||||
QY 951 TGATGGAGTCACAAAAAIAAAAAAAAAAAAAAAAAA 989

RESULT 14
LOCUS AA267728 413 bp mRNA EST 21-MAR-1997
DEFINITION mz66e03.r1 Soares mouse lymph node NbMLN Mus musculus cDNA clone
718396 5'.
ACCESSION AA267728
NID g1904462
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Mus.

REFERENCE 1 (bases 1 to 413)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HMI Mouse EST Project
Unpublished (1996)

TITLE
JOURNAL
COMMENT

Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicinep
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810



Release 2.1d John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by Intelligenetics, Inc.

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Thu Feb 26 10:19:02 1998; MasPar time 1183.69 Seconds
Tabular output not generated. 1186.804 Million cell updates/sec

Title: >US-08-874-460-1
Description: (1-989) from US08874460.seq
Perfect Score: 989
N.A. Sequence: 1 CCGGGGGGATCAGTCCTCT.....AAAAA.....
Comp: GCGCCCGGATCGAGGGA.....TTTTTTTTTTTTTTTTTTTT

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0
Searched: 430261 seqs, 710217276 bases x 2
Post-processing: Minimum Match 0%
Listing first 45 summaries

embl-new7
9:HUM2 10:HUM3 11:INV1 12:INV2 13:ORG 14:MAM 15:VRT
16:PLN 17:PRO1 18:PRO2 19:ROD 20:SYN 21:UNC 22:VIR
genbankal101
23:BC1 24:BC2 25:BC3 26:BC4 27:BC5 28:BC6 29:BC7
30:BC8 31:BC9 32:BC10 33:BC11 34:BC12 35:BC13
36:GEN1 37:GEN2 38:GEN3 39:GEN4 40:GEN5 41:GEN6 42:HTG1
43:HTG2 44:HTG3 45:HTG4 46:HTG5 47:INV1 48:INV2 49:INV3
50:INV4 51:INV5 52:INV6 53:INV7 54:INV8 55:INV9 56:INV10
57:INV11 58:INV12 59:MAM1 60:MAM2 61:MAM3 62:VRT1
63:VRT2 64:VRT3 65:VRT4 66:PAT1 67:PAT2 68:PAT3 69:PAT4
70:PAT5 71:PAT6 72:PAT7 73:PHG 74:PLN1 75:PLN2 76:PLN3
77:PLN4 78:PLN5 79:PLN6 80:PLN7 81:PLN8 82:PLN9 83:PLN10
84:PLN11 85:PLN12 86:PRI1 87:PRI2 88:PRI3 89:PRI4
90:PRI5 91:PRI6 92:PRI7 93:PRI8 94:PRI9 95:PRI10
96:PRI11 97:PRI12 98:PRI13 99:PRI14 100:PRI15 101:PRI16
102:PRI17 103:ROD1 104:ROD2 105:ROD3 106:ROD4 107:ROD5
108:ROD6 109:ROD7 110:ROD8 111:ROD9 112:STR 113:SYN
114:UNA
genbankb101
115:VRL1 116:VRL2 117:VRL3 118:VRL4 119:VRL5 120:VRL6
121:VRL7 122:VRL8 123:VRL9 124:VRL10 125:VRL11
genbank-new7
126:BC1 127:GEN 128:HTG1 129:HTG2 130:INV 131:MAM
132:VRT 133:PHG 134:PLN1 135:PLN2 136:PRI1 137:PRI2
138:ROD 139:SYN 140:UNA 141:VRL
Database: u-emb151_101
142:part1 143:part2

Statistics: Mean 11.647; Variance 10.251; scale 1.136
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description | Pred. No. | |
|------------|-------|-------------|--------|-------|-------------|------------------------|----------|
| 1 | 37 | 3.7 | 1707 | 108 | MUSIMPD | Mouse IMP dehydrogena | 5.58e-02 |
| 2 | 37 | 3.7 | 1711 | 110 | RNH10SH1 | R.norvegicus mRNA for | 5.58e-02 |
| 3 | 37 | 3.7 | 1779 | 110 | RNH10SH10 | R.norvegicus mRNA for | 5.58e-02 |
| 4 | 36 | 3.6 | 12226 | 124 | RTUPOLYP | Rice tungro spherical | 1.26e-01 |
| 5 | 35 | 3.5 | 1215 | 68 | E00200 | CDNA encoding beta-ch | 2.80e-01 |
| 6 | 35 | 3.5 | 1215 | 66 | A06805 | H.sapiens HLA-DR-beta | 2.80e-01 |
| 7 | 35 | 3.5 | 1355 | 64 | SACHORION | S.aureus mRNA for ch | 2.80e-01 |
| 8 | 35 | 3.5 | 2619 | 108 | MUSPAXIA | Mouse Pax-1 protein (| 2.80e-01 |
| 9 | 35 | 3.5 | 232650 | 136 | U60661 | Human germline T-cell | 2.80e-01 |
| 10 | 35 | 3.5 | 232650 | 10 | HSU66061 | Human germline T-cell | 2.80e-01 |
| 11 | 35 | 3.5 | 232650 | 102 | U60661 | Human germline T-cell | 2.80e-01 |
| 12 | 34 | 3.4 | 1024 | 130 | PAU69261 | Periplaneta americana | 6.20e-01 |
| 13 | 34 | 3.4 | 1194 | 110 | RNSFB | Rat sfb mRNA for sile | 6.20e-01 |
| 14 | 34 | 3.4 | 1714 | 95 | HSU63008 | Human homogenitatis d | 6.20e-01 |
| 15 | 34 | 3.4 | 1747 | 72 | I34199 | Sequence 1 from paten | 6.20e-01 |
| 16 | 34 | 3.4 | 1748 | 142 | A44314 | Sequence 1 from Paten | 6.20e-01 |
| 17 | 34 | 3.4 | 2021 | 104 | MMST6STAR | M.musculus Gal beta-1 | 6.20e-01 |
| 18 | 34 | 3.4 | 2145 | 76 | CSNADP | C.sorokiniana NADP-gd | 6.20e-01 |
| 19 | 34 | 3.4 | 2255 | 68 | E00984 | Probe detection of br | 6.20e-01 |
| 20 | 34 | 3.4 | 2255 | 70 | I04527 | Sequence 1 from Paten | 6.20e-01 |
| 21 | 34 | 3.4 | 2626 | 92 | HSAL | H.sapiens mRNA for ly | 6.20e-01 |
| 22 | 34 | 3.4 | 2850 | 107 | MUSFAUA | Mus musculus Fau gene | 6.20e-01 |
| 23 | 34 | 3.4 | 5503 | 90 | HSCH16FAA | H.sapiens mRNA for FA | 6.20e-01 |
| 24 | 33 | 3.3 | 305 | 62 | BUPBP2 | B.japonicus pBP2 mRNA | 1.36e+00 |
| 25 | 33 | 3.3 | 394 | 113 | SYNHET505 | heteropolymetric dna 5 | 1.36e+00 |
| 26 | 33 | 3.3 | 1069 | 75 | ATU60445 | Arabidopsis thaliana | 1.36e+00 |
| 27 | 33 | 3.3 | 1095 | 79 | NT1433LA | N.tabacum mRNA for 14 | 1.36e+00 |
| 28 | 33 | 3.3 | 1136 | 74 | ASU58949 | Allium sativum lectin | 1.36e+00 |
| 29 | 33 | 3.3 | 1136 | 16 | ASU58949 | Allium sativum lectin | 1.36e+00 |
| 30 | 33 | 3.3 | 1375 | 77 | EGMMDH | E.gunnii mRNA for mit | 1.36e+00 |
| 31 | 33 | 3.3 | 1596 | 56 | LTACT1596 | L.terrestris mRNA for | 1.36e+00 |
| 32 | 33 | 3.3 | 1609 | 56 | MSVATP28K | M.sextra mRNA for vacu | 1.36e+00 |
| 33 | 33 | 3.3 | 1882 | 75 | ATSUCIS | A.thaliana SUC1 mRNA | 1.36e+00 |
| 34 | 33 | 3.3 | 2334 | 79 | NEUCAM | Neurospora crassa cal | 1.36e+00 |
| 35 | 33 | 3.3 | 2526 | 65 | XHCFYN | Xiphophorus c-fyn (Xf | 1.36e+00 |
| 36 | 33 | 3.3 | 3143 | 57 | S73271 | trehalase [Bombyx mor | 1.36e+00 |
| 37 | 33 | 3.3 | 3226 | 106 | MUSAGTFCP | Mouse alpha-globin tr | 1.36e+00 |
| 38 | 33 | 3.3 | 4023 | 106 | MUSAGTFCPB | Mouse alpha-globin tr | 1.36e+00 |
| 39 | 33 | 3.3 | 6662 | 128 | PFW30005 | Plasmodium falciparum | 1.36e+00 |
| 40 | 33 | 3.3 | 32088 | 81 | SC32RBF | S.cerevisiae 32kb DNA | 1.36e+00 |
| 41 | 33 | 3.3 | 43676 | 81 | SCCEN12RG | S.cerevisiae DNA from | 1.36e+00 |
| 42 | 33 | 3.3 | 91525 | 89 | HSAF001552 | Homo sapiens chromoso | 1.36e+00 |
| 43 | 33 | 3.3 | 226841 | 46 | HSAC002043 | *** SEQUENCING IN PRO | 1.36e+00 |
| 44 | 33 | 3.3 | 226841 | 5 | HSAC2043 | *** SEQUENCING IN PRO | 1.36e+00 |
| 45 | 33 | 3.3 | 242825 | 137 | HSPEX | H.sapiens PRX gene | 1.36e+00 |

ALIGNMENTS

1
RESULT LOCUS MUSIMPD 1707 bp mRNA ROD 12-JUN-1993
DEFINITION Mouse IMP dehydrogenase mRNA, complete cds.
ACCESSION M33934
NID G198393
KEYWORDS IMP dehydrogenase.
SOURCE Mouse, cdna to mRNA.
ORGANISM Mus musculus
Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
Murinae; Mus.
REFERENCE 1 (bases 1 to 1707)
AUTHORS Tiedeman,A.A. and Smith,J.M.
TITLE Isolation and sequence of a cDNA encoding mouse IMP dehydrogenase
JOURNAL Gene 97, 289-293 (1991)
MEDLINE 91153661
COMMENT Draft entry and computer-readable sequence for [Unpublished (1990)]

[illegible]

```

Murinae; Mus.
1 (bases 1 to 2619)
Chalepakis,G., Fritsch,R., Fickenscher,H., Deutsch,U., Goulding,M.
and Gruss,P.
The molecular basis of the undulated/Pax-1 mutation
Cell 66, 873-884 (1991)
91364170

Location/Qualifiers
1..2619
/organism="Mus musculus"
/strain="C57/BL6"
/dev_stage="8.5 day embryo"
/tissue_lib="lambda gt10"
244..2535
/gene="Pax-1"
/gene="Pax-1"
244..1329
/gene="Pax-1"
/codon_start=1
/product="Pax-1 protein"
/db_xref="PID:g200224"
/translation="MQDTGEVNLGGVFNGRPLPNALRLRIVELAQLGIRPCDLSR
QRSLRGCVSKILRNKYNVESGISLPGAIGGSKPRVTTPNVKKHIRDTKQDGPFIWAEI
RORLLADKYCDIKNPFVSISIRILKNKGISLAQPGYFASKQPQPQPALPNYHIQY
PYPSPKMDTHPTGVPVGSGAHVISIPRSVSAHSVSNILGIRTFMEQTGALTGSSE
GAAYSPKMGTDAGVNRAAETPTSPAVNGLEKPALEADIKVTLASSISLVASVGGTLPACAY
PASNQGVYSAPAGCYLSLSPGWPPPAQAAPLPHFGAGVAVHGELAAAMTFKRREGTD
TCRBV20S1AIN2 genes from bases 452324 to 684973 (section 3 of 3).
2995
mutation
PAX1
/gene="Pax-1"
/note="undulated point mutation (g to a)"
polyA_signal
2530..2535
/gene="Pax-1"

BASE COUNT 558 a 805 c 660 g 596 t
ORIGIN chromosome 2.

Query Match 3.5% Score 35; DB 108; Length 2619;
Best Local Similarity 92.7%; Pred. No. 2.80e-01;
Matches 38; Conservative 0; Mismatched 3; Indels 0; Gaps 0;

Db . 2560 ggtagggagcaaaaaaaaaaaaaaaaaaaaaaa 2600
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 949 GGATGTGGAGTCATAAAAAAAAAAAAAAAAAAAAAA 989

RESULT 9
LOCUS U66061 232650 bp DNA PRI 22-JUL-1997
DEFINITION Human germline T-cell receptor beta chain TCRBV17S1ALT, TCRBV2S1,
TCRBV10S1P, TCRBV29S1P, TCRBV19S1P, TCRBV15S1, TCRBV11S1ALT, HVB
rellic, TCRBV28S1P, TCRBV34S1, TCRBV14S1, TCRBV3S1, TCRBV4S1ALT,
TRY4, TRY5, TRY6, TRY7, TRY8, TCRBD1, TCRBJ1S1, TCRBJ1S2, TCRBJ1S3,
TCRBJ1S4, TCRBJ1S5, TCRBJ1S6, TCRBCL, TCRBD2, TCRBV2S1, TCRBV2S2,
TCRBV2S3, TCRBV2S4, TCRBV2S5, TCRBV2S6, TCRBV2S7, TCRBC2,
TCRBV20S1AIN2 genes from bases 452324 to 684973 (section 3 of 3).
U66061 L36092
NID U66061
KEYWORDS C region; C-beta gene segment; D region; J-beta gene segment;
J-segment; T-cell receptor beta-chain; TCR-beta gene; V-beta gene
segment; V-segment; cell membrane protein; constant region;
diversity region; germline; joining segment; trypsin; trypsinogen;
variable segment.
Hom sapiens (clone: K41A) (clone_lib: CGM1: YAC D49H4) DNA; Homo
sap lens (clone: K35) (clone_lib: CGM1: YAC D49H4) DNA; Homo sapiens
(clone: K26) (clone_lib: CGM1: YAC D49H4) DNA; Homo sapiens (clone:
K56) (clone_lib: CGM1: YAC D49H4) DNA; Homo sapiens (library: ATCC
1521) (clone: X21B) (clone_lib: Kai Wang's) DNA; Homo sapiens
(library: ATCC 1521) (clone: G54) (clone_lib: Kai Wang's) DNA; Homo
sap lens (library: HeLa) (clone: H137) (clone_lib: Eric Lai's) DNA;
Homo sapiens (library: sperm) (clone: H18) (clone_lib: Eric Lai's)
DNA; Homo sapiens (clone: H18/G15gap) DNA; Homo sapiens (library:
ATCC 1521) (clone: G15) (clone_lib: Kai Wang's) DNA; Homo sapiens
(library: ATCC 1521) (clone: X1A) (clone_lib: Kai Wang's) DNA; Homo
sap lens (clone: A27) (clone_lib: CGM1: YAC 234 A6F6) DNA; Homo
sap lens (clone: A212partial) (clone_lib: CGM1: YAC 234 A6F6) DNA;

```

```

RT Strategy and error analysis";
RL Unpublished.
RN [3]
RP Rowen L., Wang K., Boysen C., Ahearn M.E., Charmley P., Paepers B.,
RA Lee I., Chen L., Trask B., Nickerson D., Seto D., Hood L.;
RT "Sequence variation among several haplotypes in the human T cell
RL receptor beta locus";
RN Unpublished.
RN [4]
RP Rowen L.;
RA Rowen L.;
RT Submitted (15-OCT-1994) to the EMBL/GenBank/DBJ databases.
RL L.Rowen leerowen@u.washington.edu
RN [5]
RP Rowen L.;
RA Rowen L.;
RT Submitted (25-JUN-1997) to the EMBL/GenBank/DBJ databases.
RL L.Rowen leerowen@u.washington.edu
DR IMG/LIGM; U66061; Release 97.06.
CC This sequence overlaps section 2 of the human T cell receptor beta
CC locus, Genbank Accession Number U66060, by 20616 bases.
FH Key Location/Qualifiers
FH
FT 1..232650
FT /organism="Homo sapiens"
FT /map="7q35"
FT source
FT 1..34620
FT /note="(vector pWE15A)"
FT /organism="Homo sapiens"
FT /cell_line="CGM1, haplotype A"
FT /clone="C215"
FT /clone_lib="YAC 234 A72B3"
FT /germline
FT <1..20617
FT /note="(vector pWE15A)"
FT /organism="Homo sapiens"
FT /cell_line="CGM1, haplotype A"
FT /clone="A16"
FT /clone_lib="YAC 234 A6F6"
FT /germline
FT repeat_region
FT 957..1082
FT /note="LINE; putative"
FT /rpt_family="L1MA4A"
FT repeat_region
FT 1294..1328
FT /note="SINE; putative"
FT /rpt_family="MIR"
FT join(1483..1531,1664..>1958)
FT /gene="TCRBV17S1A1T"
FT CDS
FT /note="V-segment translation product"
FT /codon_start=1
FT /db_xref="PID:g2218050"
FT /translation="MSNQVLCVCVFLGANTVDGGITQSPKYLFRKQGNVLTSCQ
FT. N
FT. P
FT LNHDAWYRQDPGGQGLRIYYISQIVNDFQKDIAGYSVSREKESFPLTVTSQAKN
FT TAFYLCASSIX"
FT join(1483..1531,1664..1958)
FT V_segment
FT /gene="TCRBV17S1A1T"
FT /standard_name="TCRBV17S1"
FT complement(2280..2579)
FT repeat_region
FT /note="putative"
FT /rpt_family="Alu"
FT 2664..2756
FT repeat_region
FT /note="(TA)46, 81% conserved; putative"
FT /rpt_type=TANDEM
FT 4775..5156
FT repeat_region
FT /note="LINE; putative"
FT /rpt_family="LINE3"
FT complement(5186..5635)
FT repeat_region
FT /note="MaLR retroposon LTR; putative"
FT
FT /rpt_family="MLT1F/G"
FT 5584..>5584
FT /note="translocation breakpoint; t(7;9); putative. Begin
FT region of similarity to chromosome 9p13."
FT 6654..6746
FT /standard_name="TCRBV relic (exon 2 only)"
FT /note="-78% similar to TCRBV7 (4) variable gene exon2
FT regions; putative"
FT 7426..7605
FT /note="DNA transposon fossil; putative"
FT /rpt_family="MER47"
FT complement(8681..8975)
FT repeat_region
FT /note="putative"
FT /rpt_family="Alu"
FT join(9153..9183,9522..9825)
FT V_segment
FT /gene="TCRBV2S1"
FT /standard_name="TCRBV2S1"
FT join(9153..9183,9522..>9825)
FT CDS
FT /gene="TCRBV2S1"
FT /note="V-segment translation product"
FT /codon_start=1
FT /db_xref="PID:g2218051"
FT /translation="MLLLLLLPGSGLGAUVSQHPSWVICKSGTSVKIECRSLDFQA
FT. T
FT. S
FT TMFWYRQFPKQSLMLMATSNESKATYEQGVKDKFLINHASLTLTLTVTSAPEDS
FT FYCSARX"
FT 10771..11307
FT /note="LINE; putative"
FT repeat_region
FT /rpt_family="L1ME"
FT 11431..11645
FT repeat_region
FT /note="DNA transposon fossil; putative"
FT /rpt_family="MER63"
FT 12880..12923
FT repeat_region
FT /note="putative"
FT /rpt_family="(AC)22"
FT /rpt_unit=465204..465205
FT complement(13511..13698)
FT repeat_region
FT /note="LINE; putative"
FT /rpt_family="MIR2"
FT 13836..14352
FT repeat_region
FT /note="Unclassified; putative"
FT /rpt_family="MER34"
FT 15731..16019
FT repeat_region
FT /note="putative"
FT /rpt_family="Alu"
FT 16056..16351
FT repeat_region
FT /note="putative"
FT /rpt_family="Alu"
FT 16432..17709
FT repeat_region
FT /note="LINE; putative"
FT /rpt_family="L1ME"
FT 17987..18281
FT repeat_region
FT /note="putative"
FT /rpt_family="Alu"
FT 18267..18303
FT repeat_region
FT /note="putative"
FT /rpt_family="(AAAT)9"
FT /rpt_unit=470593..470596
FT 18423..19033
FT repeat_region
FT /note="putative"
FT /rpt_family="LINE 1"
FT 18443..19041
FT repeat_region
FT /note="LINE; putative"
FT /rpt_family="L1PA7"
FT join(19336..19385,19499..>19796)
FT CDS
FT /gene="TCRBV10S1P"
FT /codon_start=1
FT /pseudo
FT /db_xref="PID:e325838"
FT join(19336..19385,19499..19796)
FT V_segment
FT /gene="TCRBV10S1P"
FT /standard_name="TCRBV10S1"

```


KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 1747)
AUTHORS Boulton,T.G.; Cobb,M.H.; Yancopoulos,G.D.; Nye,S. and
Panayotatos,N.
TITLE Family of map2 protein kinases
JOURNAL Patent: US 5595904-A 1 21-JAN-1997;
FEATURES Location/Qualifiers
source
1..1747
/organism="unknown"
BASE COUNT 423 a 524 c 425 g 375 t
ORIGIN

Query Match 3.4%; Score 34; DB 72; Length 1747;
Best Local Similarity 68.9%; Pred. No. 6.20e-01;
Matches 62; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Db 1650 tgcctcccccacccctgcctgtatctatataataatagagatgtatatgg 1709
|||||
QY 900 TGCCCTCCAGCTATGCTTTCTATACCTTTAATAAACCTTGGGGGTGATGGAGT 959
|||||

Db 1710 aaaaaaaaaaaaaaaaaaaaaaaaaa 1739
|||||
QY 960 CAAAAAAAAAAAAAAAAAAAAAAAAA 989
|||||

Search completed: Thu Feb 26 10:44:39 1998
Job time : 1537 secs.

W P S R L
(TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Thu Feb 26 10:45:00 1998; MasPar time 139.57 Seconds
817.716 Million cell updates/sec
Tabular output not generated.

Title: >US-08-874-460-1
Description: (1-989) from US08874460.seq
Perfect Score: 989
N.A. Sequence: 1 CCGCGCGGCATCAGTCCCT.....AAAAA.....989
Comp: GCGCGCGGCATCAGTCCCT.....TTTTTTTTTTTTTTTTTTTT

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0
Searched: 159651 seqs, 57698962 bases x 2
Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n-genseq30
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33

Statistics: Mean 9.188; Variance 6.978; scale 1.317
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

| SUMMARIES | | | | | | | |
|-----------|-------|-------|--------|----|-------------|-----------------------|----------|
| Result | Query | Match | Length | ID | Description | Pred. No. | |
| 1 | 76 | 7.7 | 1047 | 2 | Q10572 | Human Natriuretic Pep | 3.89e-24 |
| 2 | 70 | 7.1 | 1047 | 2 | Q10572 | Human Natriuretic Pep | 4.25e-21 |
| 3 | 47 | 4.8 | 204 | 1 | N81164 | Base substituted E.co | 7.76e-10 |
| 4 | 45 | 4.6 | 91 | 9 | Q31746 | Oligonucleotide probe | 6.73e-09 |
| 5 | 44 | 4.4 | 91 | 9 | Q31746 | Oligonucleotide probe | 1.97e-08 |
| 6 | 41 | 4.1 | 204 | 1 | N81164 | Base substituted E.co | 4.74e-07 |
| 7 | 39 | 3.9 | 114 | 12 | Q70465 | Generic DNA sequence | 3.84e-06 |
| 8 | 36 | 3.6 | 114 | 12 | Q70468 | Generic DNA sequence | 8.38e-05 |
| 9 | 35 | 3.5 | 114 | 12 | Q70466 | Generic DNA sequence | 2.31e-04 |
| 10 | 35 | 3.5 | 130 | 31 | T76152 | Human vascular cell a | 2.31e-04 |
| 11 | 35 | 3.5 | 190 | 31 | T76452 | Chymase antitense oli | 2.31e-04 |
| 12 | 35 | 3.5 | 1215 | 3 | N40067 | Sequence of HLA-DR-be | 2.31e-04 |
| 13 | 34 | 3.4 | 114 | 12 | Q70467 | Generic DNA sequence | 6.29e-04 |
| 14 | 34 | 3.4 | 114 | 12 | Q70469 | Generic DNA sequence | 6.29e-04 |
| 15 | 34 | 3.4 | 114 | 12 | Q70470 | Generic DNA sequence | 6.29e-04 |

| | | | | | | | |
|----|----|-----|------|----|--------|-----------------------|----------|
| 16 | 34 | 3.4 | 1538 | 29 | T63302 | A. chrysogenum O-acet | 6.29e-04 |
| 17 | 34 | 3.4 | 1738 | 10 | Q56241 | Sequence encoding ant | 6.29e-04 |
| 18 | 34 | 3.4 | 1747 | 3 | Q20260 | ERK1 | 6.29e-04 |
| 19 | 34 | 3.4 | 1748 | 15 | Q83327 | Arabidopsis thaliana | 6.29e-04 |
| 20 | 34 | 3.4 | 2255 | 2 | N60228 | Sequence of the break | 6.29e-04 |
| 21 | 33 | 3.3 | 2420 | 7 | Q46672 | p mySA cDNA clone. | 1.70e-03 |
| 22 | 33 | 3.3 | 2555 | 2 | Q11009 | Fibulin B. | 1.70e-03 |
| 23 | 33 | 3.3 | 3238 | 14 | Q83532 | Sequence encoding ser | 1.70e-03 |
| 24 | 33 | 3.3 | 3820 | 1 | Q03517 | Plasmid pFGH1 encodin | 1.70e-03 |
| 25 | 32 | 3.2 | 250 | 31 | T76438 | Substance P antisense | 4.56e-03 |
| 26 | 32 | 3.2 | 433 | 3 | Q23343 | Murine CP-10. | 4.56e-03 |
| 27 | 32 | 3.2 | 731 | 18 | Q99804 | PR-1 like gene PR-lmz | 4.56e-03 |
| 28 | 32 | 3.2 | 794 | 3 | N92416 | Sequence of the S3 CD | 4.56e-03 |
| 29 | 32 | 3.2 | 887 | 1 | Q06743 | Insert from ptIGH6 en | 4.56e-03 |
| 30 | 32 | 3.2 | 1113 | 31 | T75765 | P. tetragonolobus chi | 4.56e-03 |
| 31 | 32 | 3.2 | 1141 | 23 | T13810 | Murine VRF167 cDNA. | 4.56e-03 |
| 32 | 32 | 3.2 | 1242 | 23 | T13809 | Human nuclear prolif | 4.56e-03 |
| 33 | 32 | 3.2 | 1315 | 29 | T65022 | Tulip allylacyl amida | 4.56e-03 |
| 34 | 32 | 3.2 | 1340 | 3 | Q20497 | Human nuclear prolif | 4.56e-03 |
| 35 | 32 | 3.2 | 1393 | 6 | Q38027 | CAD cDNA isolated fro | 4.56e-03 |
| 36 | 32 | 3.2 | 1632 | 12 | Q71341 | Human basigin 1 immu | 4.56e-03 |
| 37 | 32 | 3.2 | 1718 | 23 | T40760 | Doc2 (brain-specific | 4.56e-03 |
| 38 | 32 | 3.2 | 1739 | 4 | Q27539 | Tyrosine Kinase recep | 4.56e-03 |
| 39 | 32 | 3.2 | 1741 | 22 | T32830 | Wheat adenylsuccinat | 4.56e-03 |
| 40 | 32 | 3.2 | 1831 | 15 | Q97312 | Rat A3 adenosine rece | 4.56e-03 |
| 41 | 32 | 3.2 | 2032 | 29 | T42073 | Human nuclear prolif | 4.56e-03 |
| 42 | 32 | 3.2 | 2159 | 10 | Q67223 | Mouse p55Nuc. | 4.56e-03 |
| 43 | 32 | 3.2 | 2628 | 16 | T01014 | Lupin exo-(1-4)beta-D | 4.56e-03 |
| 44 | 32 | 3.2 | 3784 | 1 | Q05326 | Sequence encoding fis | 4.56e-03 |
| 45 | 32 | 3.2 | 5894 | 33 | T85474 | hABC3 cDNA sequence. | 4.56e-03 |

ALIGNMENTS

RESULT 1
ID Q10572 standard; DNA; 1047 BP.
AC Q10572;
DT 09-APR-1991 (first entry)
DE Human Natriuretic Peptide Receptor B.
KW NPBB; ANP; BNP; CNP; kidney failure; heart failure; protein kinase;
KW hyperaldosteronism; glaucoma; guanyl cyclase.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Peptide 1..22
FT /label= signal sequence
FT Protein 12
FT /label= mature NPBB
FT Domain 23..455
FT /label= extracellular domain
FT /note= "binds natriuretic peptides A,B and C]"
FT Domain 456..456
FT /label= transmembrane domain
FT Domain 479..1047
FT /label= cytoplasmic domain
FT /note= "GC and protein kinase activity"
FT Modified-site 24..26
FT /label= N-glycos_site
FT Modified-site 35..37
FT /label= N-glycos_site
FT Modified-site 161..163
FT /label= N-glycos_site
FT Modified-site 195..197
FT /label= N-glycos_site
FT Modified-site 244..246
FT /label= N-glycos_site
FT Modified-site 277..279
FT /label= N-glycos_site
FT Modified-site 349..351
FT /label= N-glycos_site
FT Modified-site 600..602
FT /label= N-glycos_site
FN W09100292-A.
PD 10-JAN-1991.

| Matches | 77; Conservative | 249; Mismatches | 593; Indels | 7; Gaps | 7; |
|------------|---|---|-------------|---------|----|
| D b | 66 | qvrnngarrntnaavnnnnhnnsyawaivrvgnavanavnangrannvdrnrsvsnnnagc | 75 | | |
| Q y | 66 | GGGAGGAGGACCCGCCTGCAGCATGAACCTGTGGCTCCTGGCGTGCCTGTGTGGCGCGCTT | 125 | | |
| D b | 76 | snynnasavdnknphndnnnngcvynaasvarnashwrnnntntagavagsnakndh | 135 | | |
| Q y | 126 | CCTGGGACCTTGGGCCCGCCGTGTCCACACCACCAAGTGTTCTTTGAGAGACTGCTTCCTGGC | 185 | | |
| D b | 136 | yrtvrtgnsankgnvnvtvhghnhwtaraannyndartddrhnyhtnngnvnnannsgn | 195 | | |
| Q y | 186 | CTACCACTACCCATTGGGTGGGTGTGCTCCGCGCGCTGGACITACCGGATCCAGGA | 245 | | |
| D b | 196 | nsvnhvyarongnnnathanzrangrvncvgnnmhmhnnnnnrrnntngdyvnnyn | 255 | | |
| Q y | 246 | GGTGAGCGGGAGCTGCATCTCCCTGCTGGGATATTCTACCTCCCACAGACACAGGAA | 305 | | |
| D b | 256 | dvnqnsragttratgrnwndrrtnnanranntenvntyrnnnnnnnnnnnnnnnn | 315 | | |
| Q y | 306 | GGTGTGTGGGAACCCCAAAGACA - GGGAGGTGCAGAGCCATGAAGCTCTCGATGCTC | 364 | | |
| D b | 316 | rardngvynnsgnmnnnagcnydgnnnyvanvnnntnnngtgtrndgnrnvmkmgtrybhq | 375 | | |
| Q y | 365 | GAATAAGTTTTGCAAGCTCCACCACACACGCGAGACCTTCCAAGGCCCTCATGCTG | 424 | | |
| D b | 376 | vtgnvmdknrdtrtdhvwamgdn - dsqdnnnaahysganknnwtrgnnnvvkgans | 434 | | |
| Q y | 425 | TAAAGAAGTTGAGTCTGGAACATCCAAGTTATCATCGTCCAAGTTTAGCAATCCCATCA | 484 | | |
| D b | 435 | dnnncandnddscktknnstnanvangtntnmngvssnnnnrkmmnknnasmwrn | 494 | | |
| Q y | 485 | GCAGCACGACAGAGAAATGCTC - CCTCCTCATATCAAGTAATTCAGGACTGTGAGCGCGC | 543 | | |
| D b | 495 | r-wnnnnngnsryhkgagsratnszrgsygsnmtahgkyynnnaantghnkgvrvvankhv | 553 | | |
| Q y | 544 | TCATTTCTGGGCTCCATCGGCACAGGAGCGCGGATCTTTCCGATATAAACCGTCGCCCC | 603 | | |
| D b | 554 | nkkrnnntrvnnnnakhmrdrvnnhntrrngacndnnnncnvntyncrjsgnddnnndsn | 613 | | |
| Q y | 604 | TACAGACCAGCTGTCCCCAGCGCTGTCTTTTGGGTCAAGTCTTAATCCCTGCACCTG | 663 | | |
| D b | 614 | ndwmrysnndndvkgmnnhshsns - shgsnksncvvdsvrnvknktidygnasnsta | 672 | | |
| Q y | 664 | AGTTGGTCTCCCTCTGCACCCACCCACACCTCTCTGCCGTCTGGCACTGGAAAGAGGA | 723 | | |
| D b | 673 | ndddnnanyakkntannnsgnnnt-tgmnaadvysgcnnnnnnnnaat-sgmnynngdn | 731 | | |
| Q y | 724 | GTTCGCTGATTTAAGCTTTTCCGCTCCGGGACGACAGCAATCTCTGGCAGCCAG | 783 | | |
| D b | 732 | snknvkvtrngnrnyrnrsndrtnnnnnnvnmnrwan-dhaandrngnkngrnnrn | 790 | | |
| Q y | 784 | TGGCTCTGTGAGAGACATAGGATACCTCTCTCACTTTCTGTCTTCTTGGCTCCACCC | 843 | | |
| D b | 791 | kngttsndnnnnrmnyannnnknvnrntnaynnnkrkanaynnnnhsvaannkrgn | 850 | | |
| Q y | 844 | CGGGCCATGCCAGTGTCTCCCTCTGGTCCCTCCAAAACCTCTGTCAGTTCAAAGATGCC | 903 | | |
| D b | 851 | tyananadsvtnysdhvgtansanstmnnvntnnndnytcdaandnddykvtntgd | 910 | | |
| Q y | 904 | CCTCCCAGCTATCTTTTCTATAACTTTTAAATAAACCTTGGGGGTGATGGAGTCAA | 963 | | |
| D b | 911 | aymvvsngnngnrhannnarmana | 936 | | |
| Q y | 964 | AAAAAAAAAAAAAAAAAAAAA | 989 | | |
| RESULT | 3 | | | | |
| ID | N81164 | standard; DNA; 204 BP. | | | |
| AC | N81164 | | | | |
| DT | 08-NOV-1990 | (first entry) | | | |
| DE | Base substituted E.coli beta-galactosidase alpha-fragment. | | | | |
| KW | E.coli beta galactosidase alpha-fragment; base substitutions; ss. | | | | |

```

OS Escherichia coli.
FH Key Location/Qualifiers
FT misc_feature 19..69
FT /*tag_a
FT /function-multiple cloning site
FT primer_bind 187..204
FT /*tag_b
FN EP-285123-A.
PD 05-MAY-1988.
PF 30-MAR-1988; 105163.
PR 03-APR-1987; US-034819.
PA (SUSO) SUOMEN SOKERI OY.
PI Lehtovaara P, Knowles J, Koivula A, Bamford J, Reinikainen T;
DR WP; 88-279927/40.
PT Introducing random point mutations into nucleic acids -
PT by prep'n of single stranded template, annealing a primer, elongation,
PT minicorporation, completion of molecules and screening.
PS Disclosure: p; English.
CC Random point mutations were introduced into the alpha fragment of
CC E.coli beta-galactosidase. The wild type sequence was obtained as a
CC single stranded template and an oligonucleotide was hybridised to
CC it to generate a pop'n of DNA molecules which terminate at all
CC possible nucleotide positions within a specified region. The
CC variable 3' ends generated in this way are used as primers for
CC reverse transcriptase. Nucleotides are misincorporated by the
CC transcriptase and the molecules are completed to forms that can be
CC amplified and then expressed in a suitable host-vector system.
CC The sequence covers all 176 diff't base substitutions, most of which
CC occurred singularly in any given mutant.
CC See also P80575
SQ Sequence 204 BP; 21 A; 47 C; 17 G; 11 T; 108 Others;

Query Match 4.8%; Score 47; DB 1; Length 204;
Best Local Similarity 12.4%; Pred. No. 7.76e-10;
Matches 14; Conservative 59; Mismatches 40; Indels 0; Gaps 0;

Db 78 dchvgcgcymttthhyrmrbnvyrdynsdaaawccyrsvkydcocynachhdhyv 137
QY 491 GCACAGGATGTCCTCCCTCTGATATCAGTAAATTCAGGACTGTGAGCGGCTCATTTTC 550
Db 138 ybbbyvnhnnhcnccbnhvhcnvbnhrnrvyrdharrddvhhcvcvh 190
QY 551 TGGGCTCATCGGCACAGAGGCGCGGATCTTCTCCGATAAAACCGTCGCC 603

RESULT 4
ID Q51746 standard; cDNA; 91 BP.
AC Q51746;
DT 31-MAY-1994 (first entry)
DE Oligonucleotide probe MK14-A
KW Oligonucleotide; DNA probe; mycobacteria; disease diagnosis;
KW ss.
OS Synthetic.
PN EP-571911-A.
PD 01-DEC-1993.
PF 24-MAY-1993; 108325.
PR 26-MAY-1992; US-889651.
PA (BECT ) BECTON DICKINSON CO.
PI Shank DD, Spears PA;
DR WP; 93-378844/48.
PT New oligonucleotide probes specific for Mycobacteria - used for
PT detection and amplification of Mycobacteria nucleic acid in
PT samples
PS Claim 3; Page 14; 23pp; English.
CC Oligonucleotide probe MK14-A consists of nucleotides 5-95 of MK14
CC (Q51735). It hybridized to all spp. of mycobacteria tested, but
CC cross reacted to a few non-mycobacterial spp. The probe may
CC be useful as an initial screen for mycobacterial infection.
CC See also Q51735-45 and Q51747-59.
SQ Sequence 91 BP; 5 A; 17 C; 15 G; 4 T;

Query Match 4.6%; Score 45; DB 9; Length 91;
Best Local Similarity 7.3%; Pred. No. 6.73e-09;

```

[illegible]

(UUNC-) UNIV NORTH CAROLINA.
PA Fowlkes DM, Key BK;
PI WFI; 94-279739/34.
DR Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35; 255pp; English.

Q70469 is a generic DNA sequence used to generate random TSAR peptide
CC This generic formula can be represented as follows: X(TGC)(NNB)10-
CC (TGC)(NNB)62(NNB)2(TGC)(NNB)14(TGC)y. x and y are flanking restriction
CC sites (x is not the same as Y) that are not specified further. This
CC sequence generates peptides that are cloverleaf in structure. Other
CC generic sequences are shown in Q70455-68. Other specific peptides
CC generated by these generic sequences are shown in R65150-54. TSARS are
CC concatenated heterofunctional proteins or peptides, comprising at least
CC two functional regions - a binding domain with affinity for a ligand and
CC a second effector peptide portion that is chemically or biologically
CC active.They may further comprise a linker peptide between the 2 domains.
CC The oligonucleotides are also designed so that the expressed peptide
CC contains 2 or 4 cysteine residues positioned in, or flanking, the
CC unpredicted or variant residues. These residues confer some degree of
CC conformational rigidity to the peptides. The TSARS or compsns. comprising
CC a TSAR binding domain can be used in vivo to deliver a chemically or
CC biologically active moiety, eg. metal ion, radioisotope, peptide, toxin
CC or enzyme, to the specific target or on the cell. They can also replace
CC the function of macromolecules, eg. monoclonal or polyclonal antibodies
CC and therefore circumvent the need for complex methods of hybridoma
CC formation or in vivo antibody production. The TSARS are easily
CC characterised and have designated activity allowing direct and rapid
CC detection in a screening process.
SS Sequence 114 BP; 0 A; 4 G; 4 C; 4 T;

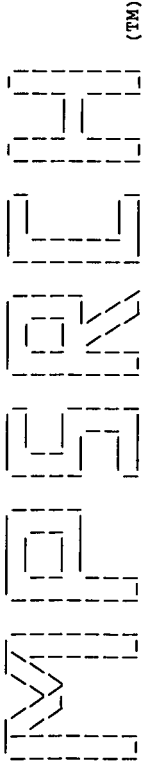
Query Match 3.4%; Score 34; DB 12; Length 114;
Best Local Similarity 5.5%; Pred. No. 6.29e-04;
Matches 6; Conservative 31; Mismatches 73; Indels 0; Gaps 0;

Dd 3 cnnbnbnbnbnbnbnbnbnbnbnbnbtgcnnbnbnbnbnbnbnbnbnbnbnnnn 62
QY | : : : : : : : : : : ||| : : : : : : : : : :
78 CGCGTCGACGATGAACCTGTGGCTCCTGCCCTGTGGTGCGCGCTCTCTGGAGCCCTG 137

Dd bnbntgcnbbnbt 112
QY :
138 GGCCCCCGCTGTCACACCACCAAGGTGTTTGAGGACTGCTCGCTGGCCCT 187

RESULT 15
ID Q70470 standard; DNA; 114 BP.
AD Q70470:
DC 10-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KK effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.

FH key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /note= "encoded by z (see comments)"
PN WO9418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994; U00977.
PP 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UUNC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Key BK;
PWFI; 94-279739/34.
DR P-PDSB; R58378.
DRI Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 36; 255pp; English.
CC Q70470 is a generic DNA sequence used to generate random TSAR (Totally



 (TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
 Distribution rights by Intelligent, Inc.

MPSrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Thu Feb 26 11:06:21 1998; MasPar time 310.45 Seconds
 Tabular output not generated. 803.702 Million cell updates/sec

Title: >US-08-874-460-1
 Description: (1-989) from US08874460.seq
 Perfect Score: 989
 N.A. Sequence: 1 CCGCGGGCAGTCAGTCCTCCT.....AAAAA.....AAAAA
 Comp: GCGCGCCGAGTCAGGGA.....TTTTTTTTTTTTTTTTTTTT

Scoring table: TABLE default

Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 333433 seqs, 126143548 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

| | | | | | | | |
|----------|----------|----------|----------|----------|----------|----------|----------|
| 1:STS1 | 2:STS2 | 3:STS3 | 4:STS4 | 5:STS5 | 6:STS6 | 7:STS7 | 8:STS8 |
| 9:STS9 | 10:STS10 | 11:STS11 | 12:STS12 | 13:STS13 | 14:STS14 | 15:STS15 | 16:STS16 |
| 17:STS17 | 18:STS18 | 19:STS19 | 20:STS20 | 21:STS21 | 22:STS22 | 23:STS23 | 24:STS24 |
| 25:STS25 | 26:STS26 | 27:STS27 | 28:STS28 | 29:STS29 | 30:STS30 | 31:STS31 | 32:STS32 |
| 33:STS33 | 34:STS34 | 35:STS35 | 36:STS36 | 37:STS37 | 38:STS38 | 39:STS39 | 40:STS40 |
| 41:STS41 | 42:STS42 | 43:STS43 | 44:STS44 | 45:STS45 | 46:STS46 | 47:STS47 | 48:STS48 |
| 49:STS49 | 50:STS50 | 51:STS51 | 52:STS52 | 53:STS53 | 54:STS54 | 55:STS55 | 56:STS56 |
| 57:STS57 | 58:STS58 | 59:STS59 | 60:STS60 | 61:STS61 | 62:STS62 | 63:STS63 | 64:STS64 |
| 65:STS65 | 66:STS66 | 67:STS67 | 68:STS68 | 69:STS69 | 70:STS70 | 71:STS71 | 72:STS72 |
| 73:STS73 | 74:STS74 | 75:STS75 | 76:STS76 | 77:STS77 | 78:STS78 | 79:STS79 | 80:STS80 |
| 81:STS81 | 82:STS82 | 83:STS83 | 84:STS84 | 85:STS85 | 86:STS86 | 87:STS87 | 88:STS88 |
| 89:STS89 | 90:STS90 | 91:STS91 | 92:STS92 | 93:STS93 | 94:STS94 | 95:STS95 | 96:STS96 |

Statistics: Mean 11.575; Variance 5.079; scale 2.279

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query Match | Length | ID | Description | Pred. No. |
|------------|-------------|--------|----|-----------------------|-----------|
| 1 | 36 | 279 | 30 | aa56b06.s1 NCI_CGAP_G | 1.59e-07 |
| 2 | 36 | 279 | 67 | HS1283434 | 1.59e-07 |

| C | 3 | 36 | 3.6 | 319 | 69 | HS1297807 | ng19g09.s1 | NCI_CGAP_L | 1.59e-07 |
|---|----|----|-----|-----|----|-----------|-----------------------|--------------|----------|
| C | 4 | 36 | 3.6 | 319 | 38 | AA501896 | ng19g09.s1 | NCI_CGAP_L | 1.59e-07 |
| | 5 | 35 | 3.5 | 221 | 42 | AA523466 | ng163h06.s1 | NCI_CGAP_P | 7.69e-07 |
| | 6 | 35 | 3.5 | 221 | 71 | HS1315081 | vh21a10.r1 | Soares mou | 7.69e-07 |
| | 7 | 35 | 3.5 | 247 | 25 | AA475511 | vh21a10.r1 | Soares mou | 7.69e-07 |
| | 8 | 35 | 3.5 | 247 | 84 | MM1266831 | aa52b07.s1 | NCI_CGAP_G | 7.69e-07 |
| C | 9 | 35 | 3.5 | 443 | 31 | AA490904 | aa52b07.s1 | NCI_CGAP_G | 7.69e-07 |
| | 10 | 35 | 3.5 | 443 | 68 | HS1288541 | aa52b07.s1 | NCI_CGAP_G | 7.69e-07 |
| C | 11 | 35 | 3.5 | 554 | 29 | AA481004 | aa28h11.s1 | NCI_CGAP_G | 7.69e-07 |
| C | 12 | 35 | 3.5 | 554 | 65 | HS1273164 | aa28h11.s1 | NCI_CGAP_G | 7.69e-07 |
| C | 13 | 35 | 3.5 | 632 | 69 | HS1299870 | aa61e05.s1 | NCI_CGAP_G | 7.69e-07 |
| C | 14 | 35 | 3.5 | 632 | 35 | AA504326 | aa61e05.s1 | NCI_CGAP_G | 7.69e-07 |
| C | 15 | 34 | 3.4 | 197 | 96 | MM1140745 | mw16e06.r1 | Soares tes | 3.65e-06 |
| C | 16 | 34 | 3.4 | 228 | 59 | HS1237988 | zw59c06.s1 | Soares tes | 3.65e-06 |
| C | 17 | 34 | 3.4 | 289 | 33 | AA496673 | zv38b11.s1 | Soares ova | 3.65e-06 |
| C | 18 | 34 | 3.4 | 269 | 69 | HS1291614 | zv38b11.s1 | Soares ova | 3.65e-06 |
| | 19 | 34 | 3.4 | 282 | 44 | C24639 | Dictyostelium discoid | 3.65e-06 | |
| | 20 | 34 | 3.4 | 282 | 46 | DDC6391 | Dictyostelium discoid | 3.65e-06 | |
| C | 21 | 34 | 3.4 | 308 | 48 | DRAA97184 | fa02e01.s1 | Zebrafish | 3.65e-06 |
| C | 22 | 34 | 3.4 | 308 | 33 | AA497184 | fa02e01.s1 | Zebrafish | 3.65e-06 |
| | 23 | 34 | 3.4 | 318 | 86 | MM1303864 | vg35f01.r1 | Soares mou | 3.65e-06 |
| | 24 | 34 | 3.4 | 318 | 36 | AA510220 | vg35f01.r1 | Soares mou | 3.65e-06 |
| C | 25 | 34 | 3.4 | 349 | 81 | MM1243962 | vc62e06.s1 | Knowles So | 3.65e-06 |
| C | 26 | 34 | 3.4 | 352 | 65 | HS1272553 | ab15h05.s1 | Stratagene | 3.65e-06 |
| C | 27 | 34 | 3.4 | 352 | 29 | AA482267 | ab15h05.s1 | Stratagene | 3.65e-06 |
| C | 28 | 34 | 3.4 | 396 | 66 | HS1279927 | aa40a10.s1 | NCI_CGAP_G | 3.65e-06 |
| C | 29 | 34 | 3.4 | 396 | 29 | AA485065 | aa40a10.s1 | NCI_CGAP_G | 3.65e-06 |
| C | 30 | 34 | 3.4 | 436 | 54 | HS1202170 | zu49f04.s1 | Soares ova | 3.65e-06 |
| C | 31 | 34 | 3.4 | 592 | 43 | AA527292 | ng39g09.s1 | NCI_CGAP_C | 3.65e-06 |
| C | 32 | 34 | 3.4 | 592 | 72 | HS1318766 | ng39g09.s1 | NCI_CGAP_C | 3.65e-06 |
| C | 33 | 33 | 3.3 | 100 | 19 | AA466951 | vf10a12.r1 | Knowles So | 1.69e-05 |
| C | 34 | 33 | 3.3 | 299 | 37 | AA512961 | nj16a11.s1 | NCI_CGAP_P | 1.69e-05 |
| C | 35 | 33 | 3.3 | 299 | 70 | HS1306663 | nj16a11.s1 | NCI_CGAP_P | 1.69e-05 |
| C | 36 | 33 | 3.3 | 307 | 58 | HS1231068 | zw77g05.s1 | Soares tes | 1.69e-05 |
| C | 37 | 33 | 3.3 | 317 | 70 | HS1301437 | n103f11.s1 | NCI_CGAP_B | 1.69e-05 |
| C | 38 | 33 | 3.3 | 337 | 33 | AA497181 | fa01h09.s1 | Zebrafish | 1.69e-05 |
| C | 39 | 33 | 3.3 | 337 | 48 | DRAA97181 | fa01h09.s1 | Zebrafish | 1.69e-05 |
| C | 40 | 33 | 3.3 | 379 | 21 | AA278864 | zs77h11.s1 | NCI_CGAP_G | 1.69e-05 |
| C | 41 | 33 | 3.3 | 418 | 48 | HS0052 | EST03894 | Homo sapiens | 1.69e-05 |
| C | 42 | 33 | 3.3 | 425 | 51 | HS1183224 | zs78d11.s1 | NCI_CGAP_G | 1.69e-05 |
| C | 43 | 33 | 3.3 | 425 | 21 | AA278817 | zs78d11.s1 | NCI_CGAP_G | 1.69e-05 |
| C | 44 | 33 | 3.3 | 519 | 72 | HS1316789 | nh36h10.s1 | NCI_CGAP_P | 1.69e-05 |
| C | 45 | 33 | 3.3 | 519 | 42 | AA525110 | nh36h10.s1 | NCI_CGAP_P | 1.69e-05 |

ALIGNMENTS

| RESULT | 1 | AA489024 | 279 bp | EST | 25-JUN-1997 |
|------------|---|--|---------------|--------------|----------------------|
| LOCUS | | aa56b06.s1 | NCI_CGAP_GCB1 | Homo sapiens | cdna clone 824915 3' |
| DEFINITION | | AA489024 | | | |
| ACCESSION | | AA489024 | | | |
| NID | | g2218626 | | | |
| KEYWORDS | | EST. | | | |
| SOURCE | | human. | | | |
| ORGANISM | | Homo sapiens | | | |
| REFERENCE | 1 | (bases 1 to 279) | | | |
| AUTHORS | | NCI-CGAP. | | | |
| TITLE | | National Cancer Institute, Cancer Genome Anatomy Project (CGAP), | | | |
| JOURNAL | | Tumor Gene Index | | | |
| COMMENT | | Unpublished (1997) | | | |

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,

Ph.D., Gerald Marti, M.D.

cdna Library Preparation: M. Bento Soares, Ph.D., M. Fatima

Bonaldo, Ph.D.

cdna Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

| | | | | | | |
|-----------------------|--|--|---------------------|---------------|-------------|-------------|
| RESULT | 5 | AA523466 | 221 bp | mRNA | EST | 17-JUL-1997 |
| LOCUS | | ni63h06.s1 NCI_CGAP_Pr12 Homo sapiens cDNA clone 981563 similar to | | | | |
| DEFINITION | | gb:X64707 BREAST BASIC CONSERVED PROTEIN 1 (HUMAN);. | | | | |
| ACCESSION | | GA223466 | | | | |
| NID | | 92264178 | | | | |
| KEYWORDS | | EST. | | | | |
| SOURCE | | human. | | | | |
| ORGANISM | | Homo sapiens | | | | |
| | | Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; | | | | |
| | | Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; | | | | |
| | | Homo. | | | | |
| REFERENCE | | 1 (bases 1 to 221) | | | | |
| AUTHORS | | NCI-CGAP. | | | | |
| TITLE | | National Cancer Institute, Cancer Genome Anatomy Project (CGAP), | | | | |
| JOURNAL | | Tumor Gene Index | | | | |
| COMMENT | | Unpublished (1997) | | | | |
| | | Contact: Robert Strausberg, Ph.D. | | | | |
| | | Tel: (301) 496-1550 | | | | |
| | | Email: Robert.Strausberg@nih.gov | | | | |
| | | Tissue Procurement: W Douglas Figg, Ph.D., Paul H. Duray, M.D., | | | | |
| | | Rodrigo F. Chuauqui, M.D., Michael R. Emmert-Buck, M.D., Ph.D. | | | | |
| | | cDNA Library Preparation: David B. Krizman, Ph.D. | | | | |
| | | cDNA Library Arrayed by: Greg Lennon, Ph.D. | | | | |
| | | DNA Sequencing by: Washington University Genome Sequencing Center | | | | |
| | | Clone distribution: NCI-CGAP clone distribution information can be | | | | |
| | | found through the I.M.A.G.E. Consortium/LLNL at: | | | | |
| | | www.bio.llnl.gov/dbrrp/image/image.html | | | | |
| | | Seq primer: -40ml3 fwd. ET from Amersham | | | | |
| | | High quality sequence stop: 171. | | | | |
| FEATURES | | Location/Qualifiers | | | | |
| source | | 1..221 | | | | |
| | | /organism="Homo sapiens" | | | | |
| | | /note="Vector: pAMP10; mRNA made from metastatic prostate | | | | |
| | | lesion of the bone, cDNA made by oligo-dT priming. | | | | |
| | | Non-directionally cloned. Size-selected on agarose gel, | | | | |
| | | average insert size 600 bp. Library made by D. Krizman, | | | | |
| | | NH. | | | | |
| | | /clone="981563" | | | | |
| | | /clone_lib="NCI_CGAP_Pr12" | | | | |
| | | /sex="male" | | | | |
| | | /tissue.type="metastatic prostate bone lesion" | | | | |
| | | /lab_host="DH10B" | | | | |
| mrna | | <1..>221 | | | | |
| BASE COUNT | | 93 a 46 c 52 g 30 t | | | | |
| ORIGIN | | | | | | |
| Query Match | | 3.5% | Score 35; | DB 42; | Length 221; | |
| Best Local Similarity | | 90.7%; | Pred. No. 7.69e-07; | | | |
| Matches | | 39; | Conservative 0; | Mismatches 4; | Indels 0; | Gaps 0; |
| Db | 162 | ggggacttggaatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 204 | | | | |
| | | | | | | |
| QY | 947 | GGGGTGTGGAGTCACAAAAAATAAAAAAAAAAAAAAAAAA 989 | | | | |
| RESULT | 6 | | | | | |
| ID | HS1315081 | standard; RNA; EST; 221 BP. | | | | |
| AC | AA523466; | | | | | |
| NI | 92264178 | | | | | |
| DT | 18-JUL-1997 (Rel. 52, Created) | | | | | |
| DT | 18-JUL-1997 (Rel. 52, Last updated, Version 1) | | | | | |
| DE | ni63h06.s1 NCI_CGAP_Pr12 Homo sapiens cDNA clone 981563 similar to | | | | | |
| DE | gb:X64707 BREAST BASIC CONSERVED PROTEIN 1 (HUMAN);. | | | | | |
| DE | EST. | | | | | |
| KW | Homo sapiens (human) | | | | | |
| OS | Homo sapiens | | | | | |
| OC | Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; | | | | | |
| OC | Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; | | | | | |
| OC | Homo. | | | | | |

[illegible]

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

Run on: Thu Feb 26 10:16:16 1998; MasPar time 6.46 Seconds
Tabular output not generated.

Title: >US-08-874-460-2
Description: (1-149) from US08874460.pep
Perfect Score: 1099
Sequence: 1 MNWLACINAGFLGAWAPA.....NPISRRNVSLLISANSGL 149

Scoring table: PAM 150
Gap 11

Searched: 111726 seqs, 13889129 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq30

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23

Statistics: Mean 31.118; Variance 123.942; scale 0.251

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description | Pred. No. |
|------------|-------|-------------|--------|----|-----------------------|-----------|
| 1 | 157 | 14.3 | 134 | 17 | Chemokine beta-9 | 2.19e-05 |
| 2 | 152 | 13.8 | 134 | 20 | Pancreas expressed ch | 6.20e-05 |
| 3 | 137 | 12.5 | 93 | 23 | Mouse thymus and acti | 1.35e-03 |
| 4 | 136 | 12.4 | 79 | 23 | Stem cell mobilising | 1.65e-03 |
| 5 | 126 | 11.5 | 94 | 23 | Human thymus and acti | 1.23e-02 |
| 6 | 125 | 11.4 | 75 | 23 | Human chemokine beta- | 1.50e-02 |
| 7 | 125 | 11.4 | 76 | 23 | Human chemokine beta- | 1.50e-02 |
| 8 | 125 | 11.4 | 77 | 23 | Human chemokine beta- | 1.50e-02 |
| 9 | 125 | 11.4 | 82 | 23 | Stem cell mobilising | 1.50e-02 |
| 10 | 125 | 11.4 | 82 | 23 | Human chemokine beta- | 1.50e-02 |
| 11 | 125 | 11.4 | 120 | 21 | Human cytokine beta-8 | 1.50e-02 |
| 12 | 125 | 11.4 | 121 | 13 | Macrophage Inflammato | 1.50e-02 |
| 13 | 125 | 11.4 | 137 | 22 | Human eosinophil-expr | 1.50e-02 |
| 14 | 119 | 10.8 | 97 | 23 | Human eosinocyte CC t | 4.91e-02 |
| 15 | 119 | 10.8 | 97 | 22 | Human eotaxin. | 4.91e-02 |
| 16 | 118 | 10.7 | 96 | 17 | Human chemokine beta- | 5.98e-02 |
| 17 | 118 | 10.7 | 96 | 18 | Liver expressed chemo | 5.98e-02 |
| 18 | 117 | 10.6 | 97 | 20 | Pancreas expressed ch | 7.27e-02 |
| 19 | 113 | 10.3 | 69 | 7 | LD78 Glu55>Arg, Glu56 | 1.58e-01 |
| 20 | 113 | 10.3 | 70 | 23 | Stem cell mobilising | 1.58e-01 |

| | | | | | | | |
|----|-----|------|-----|----|--------|-----------------------|----------|
| 21 | 113 | 10.3 | 119 | 16 | R85779 | Human monocyte chemot | 1.58e-01 |
| 22 | 113 | 10.3 | 119 | 21 | W07845 | Human monocyte chemot | 1.58e-01 |
| 23 | 112 | 10.2 | 69 | 7 | R39123 | LD78 Trp57>Ala. | 1.92e-01 |
| 24 | 112 | 10.2 | 96 | 23 | W14991 | Guinea pig eosinocyte | 1.92e-01 |
| 25 | 111 | 10.1 | 69 | 7 | R39090 | LD78 Tyr61>Ala. | 2.32e-01 |
| 26 | 108 | 9.8 | 69 | 7 | R39089 | LD78 Gln33>Ser. | 4.13e-01 |
| 27 | 108 | 9.8 | 70 | 23 | W17660 | Stem cell mobilising | 4.13e-01 |
| 28 | 107 | 9.7 | 69 | 7 | R39127 | LD78 Tyr61>ASP. | 4.99e-01 |
| 29 | 107 | 9.7 | 69 | 7 | R39125 | LD78 Trp57>Leu. | 4.99e-01 |
| 30 | 106 | 9.6 | 69 | 7 | R39106 | LD78 Val39>Ala. | 6.04e-01 |
| 31 | 106 | 9.6 | 69 | 7 | R38980 | LD78 Glu55>Ser. | 6.04e-01 |
| 32 | 106 | 9.6 | 69 | 7 | R38973 | LD78 Leu67>Ala. | 6.04e-01 |
| 33 | 105 | 9.6 | 69 | 7 | R39105 | LD78 Gly38>Ala. | 7.30e-01 |
| 34 | 105 | 9.6 | 69 | 7 | R38938 | LD78 Phe28>Glu. | 7.30e-01 |
| 35 | 104 | 9.5 | 69 | 7 | R38963 | LD78 Ser46>Glu. | 8.81e-01 |
| 36 | 104 | 9.5 | 69 | 7 | R39137 | LD78 Glu55>Gln, Glu56 | 8.81e-01 |
| 37 | 104 | 9.5 | 69 | 7 | R39132 | LD78 Asp64>Arg. | 8.81e-01 |
| 38 | 104 | 9.5 | 69 | 7 | R38940 | LD78 Phe28>Glu, Gln48 | 8.81e-01 |
| 39 | 104 | 9.5 | 69 | 7 | R38983 | LD78 Asp64>Ser. | 8.81e-01 |
| 40 | 103 | 9.4 | 69 | 7 | R38981 | LD78 Glu56>Ser. | 1.06e+00 |
| 41 | 103 | 9.4 | 69 | 7 | R39109 | LD78 Thr43>Ala. | 1.06e+00 |
| 42 | 103 | 9.4 | 92 | 20 | W01805 | Murine macrophage-der | 1.06e+00 |
| 43 | 102 | 9.3 | 69 | 7 | R39108 | LD78 Gln21>Ser. | 1.28e+00 |
| 44 | 102 | 9.3 | 69 | 7 | R38960 | LD78 Gln18>Ser. | 1.28e+00 |
| 45 | 102 | 9.3 | 73 | 13 | R70252 | Eotaxin chemoattracta | 1.28e+00 |

ALIGNMENTS

RESULT 1
ID R81567 standard; Protein; 134 AA.
AC R81567;
DT 01-OCT-1996 (first entry)
DE Chemokine beta-9.
KW Human; chemokine beta-9; Ck beta-9; breast lymph node; chemokine;
KW eotaxin; inhibition; bone marrow stem cell colony formation;
KW cancer chemotherapy; leukaemia; epidermal keratinocyte proliferation;
KW psoriasis; solid tumour; autoimmune disease; IL-2 biosynthesis.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Peptide 1..23
FT /note= "Leader sequence"
FT Protein 24..134
FT /note= "Mature Ck beta-9"
PN W09606169-Al.
PD 29-FEB-1996.
PF 06-JUN-1995; U06260.
PR 23-AUG-1994; US-294251.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Adams MD, Li H;
DR WPI: 96-151372/15.
DR N-PSDB; T18015.
PT Human chemokine beta-9 polynucleotide(s), polypeptide(s) and
PT antagonists - useful in treatment of e.g. leukaemia, tumours,
PT chronic infections and auto-immune disorders and diagnosis of Ck
PT beta-9 mutation(s)
PS Claim 12; Page 41; 55pp; English.
CC This sequence represents human chemokine beta-9 (Ck beta-9). The
CC polynucleotide encoding Ck beta-9 was discovered in a cDNA library
CC derived from a human breast lymph node. Ck beta-9 is structurally
CC related to the chemokine family. It exhibits the highest degree of
CC homology to eotaxin with 32% identity and 69% similarity over a
CC stretch of 75 amino acids. The four spatially conserved cysteine
CC residues found in chemokines are also found in Ck beta-9. Ck beta-9
CC polypeptides may be used to inhibit bone marrow stem cell colony
CC formation as adjunct protective treatment during cancer chemotherapy and
CC for leukaemia. They can also be used to inhibit epidermal keratinocyte
CC proliferation for treatment of psoriasis. They may be used to treat
CC solid tumours by stimulating the invasion and activation of host defence
CC cells, e.g. cytotoxic T cells and macrophages. They act to enhance host
CC defences against resistant chronic infections, and to treat autoimmune
CC disease and lymphocytic leukaemia by inhibiting T cell proliferation by
CC the inhibition of IL-2 biosynthesis.

08-MAR-1994; US-208339.
(HUMA-) HUMAN GENOME SCI INC.
Adams MD, Li H, Rosen CA, Ruben S;
WPI: 95-240404/31.
N-PSDB: Q94092.

CNA encoding human macrophage inflammatory proteins -3, -4 and
-1-gamma - used in immuno-regulation including inflammatory
activity, haematopoiesis, treatment of psoriasis or solid tumours.
Claim 10; Figure 1; 60pp; English.

CNA Human macrophage inflammatory protein-3 (MIP-3) is used in
therapeutic and diagnostic applications for detecting and
treating infections, cancer, inflammation, myelopoietic
dysfunction and autoimmune diseases. Antagonists/inhibitors
of MIP-3 are used to treat diseases involving overexpression
of MIP-3, including silicosis, arteriosclerosis, autoimmune and
chronic inflammatory and infective diseases, aplastic anaemia,
etc.

Sequence 121 AA;

Query Match 11.4%; Score 125; DB 13; Length 121;
Best Local Similarity 42.2%; Pred. No. 1.50e-02;
Matches 19; Conservative 11; Mismatches 12; Indels 3; Gaps 3;

Ddb 74 etnsecskpgvif-ltkkgrfcanpsdkqvqvcmrmlkdtrik 117
|::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 53 EYSGSNCNLPAAIFFLPKRHRKVGCGNPKSREVRAMK-L-LDARNK 95

RESULT 13

ID W05186 standard; Protein; 137 AA.

AC W05186;

DE 13-AUG-1997 (first entry)

KW Human eosinophil-expressed chemokine protein sequence.

DD Human eosinophil-expressed chemokine; Hypereosinophilic Syndrome;
homology; antibody; antisense; inhibitor; eosinophilia; inflammation;
viral; bacterial; fungal; parasitic; infection; mechanical injury;
trauma; hereditary allergy; asthma; leukaemia; lymphoma.

OS Homo sapiens.

PN W09632481-A1.

PD 17-OCT-1996.

PF 12-APR-1996; U05102.

PP 13-APR-1995; US-421144.

PR (INCY-) INCYTE PHARM INC.

PPI Bandman O, Braxton SM, Cocks B, Coleman R, Rhodes ET;
Stuart SG;

DR WPI: 96-477132/47.

DR N-PSDB: T43001.

DR New human eosinophil-expressed chemokine - used for producing
inhibitors for treatment of e.g. eosinophilia, inflammation,
allergies, asthma, leukaemia and lymphoma

PPT Claim 9; Page 35; 68pp; English.

CCC This is the amino acid sequence of a novel human eosinophil-expressed
chemokine (EEC). The gene sequence was isolated from human eosinophil
cDNA library constructed from cDNA generated from mRNA purified from
eosinophils from a male caucasian diagnosed with Hypereosinophilic
Syndrome. The cDNA clones obtained were sequenced and encoded proteins
were searched for homology with databases. This sequence was found to
be homologous but different from other C-C chemokine family members.
EEC-specific antibodies, antisense oligonucleotides and other inhibitors
are useful for the treatment of eosinophilia, inflammation or any disease
involving an altered number of eosinophils, e.g., viral, bacterial,
fungal or parasitic infection; mechanical injury associated with trauma;
hereditary allergies and asthma; invasive disease such as leukaemia or
lymphoma.

CC Sequence 137 AA;

CC Query Match 11.4%; Score 125; DB 22; Length 137;
CC Best Local Similarity 42.2%; Pred. No. 1.50e-02;
CC Matches 19; Conservative 11; Mismatches 12; Indels 3; Gaps 3;

Ddb 90 etnsecskpgvif-ltkkgrfcanpsdkqvqvcmrmlkdtrik 133
|::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 53 EYSGSNCNLPAAIFFLPKRHRKVGCGNPKSREVRAMK-L-LDARNK 95

Search completed: Thu Feb 26 10:16:47 1998
Job time : 31 secs.

(TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

```
Run on: Thu Feb 26 10:17:06 1998; MasPar time 8.84 Seconds
513.408 Million cell updates/sec
```

Title: >US-08-874-460-2
Description: (1-149) from US08874460.pep

Scoring table: PAM 150
Gap 11

Searched: 95051 seqs, 30469580 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

```
Database:
pir53
1:ann1 2:ann2 3:ann3 4:ann4 5:unann1 6:unann2 7:unann3
8:unann4 9:unann5 10:unann6 11:unann7 12:unann8
13:unann9 14:unann10 15:unann11 16:unann12 17:unann13
18:unrev
```

Statistics: Mean 41.426; Variance 82.134; scale 0.504

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

| Result | Query | Score | Match | Length | ID | Description | Pred. No. |
|--------|-------|-------|-------|--------|----|-------------|-----------|
| 1 | 1 | 100 | 100 | 100 | 1 | 1 | 1 |
| 2 | 2 | 100 | 100 | 100 | 2 | 2 | 2 |
| 3 | 3 | 100 | 100 | 100 | 3 | 3 | 3 |
| 4 | 4 | 100 | 100 | 100 | 4 | 4 | 4 |
| 5 | 5 | 100 | 100 | 100 | 5 | 5 | 5 |
| 6 | 6 | 100 | 100 | 100 | 6 | 6 | 6 |
| 7 | 7 | 100 | 100 | 100 | 7 | 7 | 7 |
| 8 | 8 | 100 | 100 | 100 | 8 | 8 | 8 |
| 9 | 9 | 100 | 100 | 100 | 9 | 9 | 9 |
| 10 | 10 | 100 | 100 | 100 | 10 | 10 | 10 |
| 11 | 11 | 100 | 100 | 100 | 11 | 11 | 11 |
| 12 | 12 | 100 | 100 | 100 | 12 | 12 | 12 |
| 13 | 13 | 100 | 100 | 100 | 13 | 13 | 13 |
| 14 | 14 | 100 | 100 | 100 | 14 | 14 | 14 |
| 15 | 15 | 100 | 100 | 100 | 15 | 15 | 15 |
| 16 | 16 | 100 | 100 | 100 | 16 | 16 | 16 |
| 17 | 17 | 100 | 100 | 100 | 17 | 17 | 17 |
| 18 | 18 | 100 | 100 | 100 | 18 | 18 | 18 |
| 19 | 19 | 100 | 100 | 100 | 19 | 19 | 19 |
| 20 | 20 | 100 | 100 | 100 | 20 | 20 | 20 |
| 21 | 21 | 100 | 100 | 100 | 21 | 21 | 21 |
| 22 | 22 | 100 | 100 | 100 | 22 | 22 | 22 |
| 23 | 23 | 100 | 100 | 100 | 23 | 23 | 23 |
| 24 | 24 | 100 | 100 | 100 | 24 | 24 | 24 |
| 25 | 25 | 100 | 100 | 100 | 25 | 25 | 25 |
| 26 | 26 | 100 | 100 | 100 | 26 | 26 | 26 |
| 27 | 27 | 100 | 100 | 100 | 27 | 27 | 27 |
| 28 | 28 | 100 | 100 | 100 | 28 | 28 | 28 |
| 29 | 29 | 100 | 100 | 100 | 29 | 29 | 29 |
| 30 | 30 | 100 | 100 | 100 | 30 | 30 | 30 |
| 31 | 31 | 100 | 100 | 100 | 31 | 31 | 31 |
| 32 | 32 | 100 | 100 | 100 | 32 | 32 | 32 |
| 33 | 33 | 100 | 100 | 100 | 33 | 33 | 33 |
| 34 | 34 | 100 | 100 | 100 | 34 | 34 | 34 |
| 35 | 35 | 100 | 100 | 100 | 35 | 35 | 35 |
| 36 | 36 | 100 | 100 | 100 | 36 | 36 | 36 |
| 37 | 37 | 100 | 100 | 100 | 37 | 37 | 37 |
| 38 | 38 | 100 | 100 | 100 | 38 | 38 | 38 |
| 39 | 39 | 100 | 100 | 100 | 39 | 39 | 39 |
| 40 | 40 | 100 | 100 | 100 | 40 | 40 | 40 |
| 41 | 41 | 100 | 100 | 100 | 41 | 41 | 41 |
| 42 | 42 | 100 | 100 | 100 | 42 | 42 | 42 |
| 43 | 43 | 100 | 100 | 100 | 43 | 43 | 43 |
| 44 | 44 | 100 | 100 | 100 | 44 | 44 | 44 |
| 45 | 45 | 100 | 100 | 100 | 45 | 45 | 45 |
| 46 | 46 | 100 | 100 | 100 | 46 | 46 | 46 |
| 47 | 47 | 100 | 100 | 100 | 47 | 47 | 47 |
| 48 | 48 | 100 | 100 | 100 | 48 | 48 | 48 |
| 49 | 49 | 100 | 100 | 100 | 49 | 49 | 49 |
| 50 | 50 | 100 | 100 | 100 | 50 | 50 | 50 |
| 51 | 51 | 100 | 100 | 100 | 51 | 51 | 51 |
| 52 | 52 | 100 | 100 | 100 | 52 | 52 | 52 |
| 53 | 53 | 100 | 100 | 100 | 53 | 53 | 53 |
| 54 | 54 | 100 | 100 | 100 | 54 | 54 | 54 |
| 55 | 55 | 100 | 100 | 100 | 55 | 55 | 55 |

| | | | | | | | |
|----|-----|------|-----|----|--------|-----------------------|----------|
| 1 | 122 | 11.1 | 97 | 6 | JC4912 | eotaxin - human | 5.82e-05 |
| 2 | 116 | 10.6 | 116 | 16 | I49555 | gene C10 protein - m | 4.50e-04 |
| 3 | 112 | 10.2 | 96 | 6 | I48039 | eotaxin precursor - | 1.71e-03 |
| 4 | 105 | 9.6 | 92 | 6 | C30552 | macrophage inflammat | 1.68e-02 |
| 5 | 106 | 9.6 | 96 | 6 | JC2478 | eotaxin - rat | 1.22e-02 |
| 6 | 102 | 9.3 | 92 | 6 | C30574 | T-cell activation pr | 4.35e-02 |
| 7 | 102 | 9.3 | 99 | 6 | JC2417 | monocyte chemotattrac | 4.35e-02 |
| 8 | 101 | 9.2 | 148 | 6 | C30209 | PDGF-inducible JE g1 | 5.95e-02 |
| 9 | 100 | 9.1 | 93 | 6 | B35673 | LD78-beta protein pr | 8.13e-02 |
| 10 | 98 | 8.9 | 92 | 6 | A32393 | macrophage inflammat | 1.51e-01 |
| 11 | 97 | 8.8 | 246 | 3 | R5D07 | ribosomal protein L7 | 2.05e-01 |
| 12 | 97 | 8.6 | 476 | 16 | JC4646 | bone morphogenetic p | 2.05e-01 |
| 13 | 95 | 8.6 | 571 | 10 | S52750 | dehRI protein - pseu | 3.77e-01 |
| 14 | 92 | 8.4 | 91 | 6 | A28815 | monocyte chemotattrac | 9.23e-01 |
| 15 | 91 | 8.3 | 109 | 6 | A54678 | monocyte chemotactic | 1.24e+00 |
| 16 | 91 | 8.3 | 341 | 10 | S51265 | trsf protein - Yers1 | 1.24e+00 |
| 17 | 91 | 8.3 | 534 | 9 | S55358 | DNA helicase-primase | 1.24e+00 |
| 18 | 91 | 8.3 | 847 | 10 | S35418 | mdoH protein - Esche | 1.24e+00 |
| 19 | 90 | 8.2 | 114 | 3 | ETMSL | lymphotactin precurs | 1.66e+00 |
| 20 | 88 | 8.2 | 288 | 15 | A45803 | B-cell-restricted an | 2.95e+00 |

[illegible]


```

##cross-references GB:J04491
REFERENCE - S04533
#authors Davatelis, G.; Tekamp-Olson, P.; Wolpe, S.D.; Hermesen, K.;
Luedke, C.; Gallegos, C.; Coit, D.; Merryweather, J.;
Cerami, A.
#journal J. Exp. Med. (1988) 167:1939-1944
#title Cloning and characterization of a cDNA for murine macrophage
inflammatory protein (MIP), a novel monokine with
inflammatory and chemokinetic properties.
#cross-references MUID:88258380
#accession S04533
##molecule_type mRNA
##residues 1-48,'E',50-90,'I',92 ##label DA2
##cross-references EMBL:X12531
##note the authors translated the codon GAG for residue 49 as
Asp and ATT for residue 91 as Asn
the sequence has been corrected in reference A53885
##note
#authors Davatelis, G.; Tekamp-Olson, P.; Wolpe, S.D.; Hermesen, K.;
Luedke, C.; Gallegos, C.; Coit, D.; Merryweather, J.;
Cerami, A.
#journal J. Exp. Med. (1989) 170:2189
#contents erratum
#accession A53885
##molecule_type mRNA
##residues 1-92 ##label DAV
##cross-references EMBL:X12531
REFERENCE A30552
#authors Brown, K.D.; Zurawski, S.M.; Mosmann, T.R.; Zurawski, G.
J. Immunol. (1989) 142:679-687
#journal A family of small inducible proteins secreted by leukocytes
are members of a new superfamily that includes leukocyte
and fibroblast-derived inflammatory agents, growth factors,
and indicators of various activation processes.
#cross-references MUID:89093958
#accession A30552
##molecule_type mRNA
##residues 1-21,'L',23-61,'A',63-92 ##label BRO
##cross-references GB:X23447
REFERENCE J0088
#authors Sherry, B.; Tekamp-Olson, P.; Gallegos, C.; Bauer, D.;
Davatelis, G.; Wolpe, S.D.; Maslarsz, F.; Coit, D.; Cerami,
A.
#journal J. Exp. Med. (1988) 168:2251-2259
#title Resolution of the two components of macrophage inflammatory
protein 1, and cloning and characterization of one of those
components, macrophage inflammatory protein 1 beta.
#cross-references MUID:89067830
#accession P50303
##molecule_type mRNA
##residues 24-33,'XX',36-54 ##label SHE
REFERENCE A27596
#authors Wolpe, S.D.; Davatelis, G.; Sherry, B.; Beutler, B.; Hesse,
D.G.; Nguyen, H.T.; Moldawer, L.L.; Nathan, C.F.; Lowry,
S.F.; Cerami, A.
#journal J. Exp. Med. (1988) 167:570-581
#title Macrophages secrete a novel heparin-binding protein with
inflammatory and neutrophil chemokinetic properties.
#cross-references MUID:88154745
#accession A27596
##molecule_type protein
##residues 24-33,'XX',36-42 ##label WOL
##note 26-Met, 30-Pro, and 39-Thr were also found
REFERENCE I56104
#authors Widmer, U.; Yang, Z.; van Derventer, S.; Manogue, K.R.;
Sherry, B.; Cerami, A.
#journal J. Immunol. (1991) 146:4031-4040
#title Genomic structure of murine macrophage inflammatory
protein-1-alpha and conservation of potential regulatory
sequences with a human homolog, Lb78.
#cross-references MUID:91237116
#accession I56104
##status preliminary; translated from GB/EMBL/DBJ

```

```

##molecule_type DNA
##residues 1-92 ##label RES
##cross-references GB:M73061; NID:g199694; CDS_PID:g199695
COMMENT This protein is a monokine.
GENETICS
#introns 23/3; 26/1; 63/2
CLASSIFICATION #superfamily macrophage inflammatory protein
KEYWORDS heparin binding
FEATURE
1-23 #domain signal sequence #status predicted #label SIG\
24-92 #product macrophage inflammatory protein #status
experimental #label MAT
SUMMARY #length 92 #molecular-weight 10345 #checksum 5009
Query Match 8.9%; Score 98; DB 6; Length 92;
Best Local Similarity 44.4%; Pred. No. 1.51e-01;
Matches 12; Conservative 8; Mismatches 0; Indels 1; Gaps 1;
Db 52 eteslcsqpgvif-ltkrnrgicadsk 77
| | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | |
QY 53 EVSGSCNLPAAIFVLPKRHRKVCNPK 79
| | | | | | | | | | | | | | | | | | | | | |
RESULT 11
ENTRY R5D07 #type complete
TITLE ribosomal protein L7 - slime mold (Dictyostelium discoideum)
ORGANISM #formal_name Dictyostelium discoideum
DATE 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change
30-Jun-1993
ACCESSIONS S04849; S12834
REFERENCE S04849
#authors Szymkowski, D.E.; Kelly, B.; Deering, R.A.
#journal Nucleic Acids Res. (1989) 17:5393
#title A Dictyostelium discoideum cDNA coding for a protein with
homology to the rat ribosomal protein L7.
#cross-references MUID:89345108
#accession S04849
##molecule_type mRNA
##residues 1-246 ##label SZY
##cross-references EMBL:X14909
GENETICS
#gene rpl7
CLASSIFICATION #superfamily rat ribosomal protein L7
KEYWORDS protein biosynthesis; ribosome
SUMMARY #length 246 #molecular-weight 27932 #checksum 1533
Query Match 8.8%; Score 97; DB 3; Length 246;
Best Local Similarity 36.8%; Pred. No. 2.05e-01;
Matches 21; Conservative 12; Mismatches 22; Indels 2; Gaps 2;
Db 71 riaknsgtfvypaakvafvirgngvskprkviklrlqlnggvfklnkas 127
| | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | |
QY 50 RIQEVSSCNLP-AAIFVLPKRHRKVCNPKSREVQBAKLLDARKVKFALHNT 104
| | | | | | | | | | | | | | | | | | | | | |
RESULT 12
ENTRY JC4646 #type complete
TITLE bone morphogenetic protein-3b precursor - rat
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE 10-May-1996 #sequence_revision 19-Jul-1996 #text_change
19-Jul-1996
ACCESSIONS JC4646
REFERENCE JC4646
#authors Takao, M.; Hino, J.; Takeshita, N.; Konno, Y.; Nishizawa, T.;
Matsuo, H.; Kangawa, K.
#journal Biochem. Biophys. Res. Commun. (1996) 219:656-662
#title Identification of rat bone morphogenetic protein-3b (BMP-3b),
a new member of BMP-3.
#accession JC4646
##molecule_type mRNA
##residues 1-476 ##label TAK
##cross-references DDBJ:D49494
##experimental_source femur

```

```

##residues      1-109 ##label MIN
##cross-references EMBL:X71087
COMMENT  This protein induces proteinase secretion and chemotaxis by
          macrophages and monocytes.
GENETICS
#gene          GDB:SCYA7; SCYA6; MCP-3
##cross-references GDB:138473
#map_position  17q11-17q12
#introns       36/1; 75/2
CLASSIFICATION #superfamily macrophage inflammatory protein
KEYWORDS       cytokine; glycoprotein; inflammation
FEATURE
1-33           #domain signal sequence #status predicted #label sig\
34-109         #product monocyte chemotactic protein 3 #status
39             #predicted #label MAT\
              #binding_site carbohydrate (Asn) (covalent) #status
              Predicted
SUMMARY        #length 109 #molecular-weight 12356 #checksum 1535

Query Match      8.3%; Score 91; DB 6; Length 109;
Best Local Similarity 29.2%; Pred.No. 1.24e+00;
Matches 21; Conservative 17; Mismatches 27; Indels 7; Gaps 6;

Db  44 cc-yrf-innkipqrlesyr-rttsshcp-reavifktkldeicadptqkvvgdfmk 98
    || |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
Qy  30 CCLAYHYPIGWAVLR-RAWTYRIQEVSGSCNLPAAIFLFPKRHRKVCGNPKSREVQRA 88
    || |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
Db  99 hldkktqt-pkl 109
    || |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
Qy  89 LLDARNKVFAL 100
    || |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|

```

Search completed: Thu Feb 26 10:17:38 1998
Job time : 32 secs.

Moreno-Palancas, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M., Phillips, C.A., Ryder, S.E., Scott, J.L., Saudek, D.M., Shirley, R., Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y., Bednarik, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J., Dimke, D., Feng, D.-F., Ferrie, A., Fischer, C., Hastings, G.A., He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Kozak, D.L., Kunsch, C., Hungjun, J., Li, H., Weissner, P.S., Olsen, H., Raymond, L., Wei, Y.F., Wing, J., Xu, C., Yu, G.L., Ruben, S.M., Dillion, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M. and Venter, J.C.

TITLE Initial assessment-of-human-gene-diversity and expression patterns based upon 83 million nucleotides of cDNA sequence

JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)

MEDLINE 9602680

COMMENT Other_ESTs: THC148619

Contact: Kerlavage, AR

Bioinformatics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056

Fax: 3018699423

Email: arkerlav@tigr.org

For clone availability, additional sequence and expression

information related to this EST, please check the TIGR Human Gene

Index (<http://www.tigr.org/tdb/hgi/html>)

Seq primer: M13 Reverse.

FEATURES Location/Qualifiers

source

1..291

/organism="Homo sapiens"

/note="Organ: thymus; Vector: pBluescript SK-; Site_1:

ECORI; Site_2: XhoI"

/clone_lib="Thymus III"

/dev_stage="adult"

<1..>291

BASE COUNT 48 a 98 c 64 g 76 t 5 others

ORIGIN

Query Match 78.9%; Score 286; DB 25; Length 291;

Best Local Similarity 98.3%; Pred. No. 0.00e+00;

Matches 286; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Db 1 tctttctcgataaacctgcctacagaccagctgtcccaacgctctntttttg 60

Qy 579 TCTTTCGATAAACCGTCGCTACAGACCCAGCTGCCACGCTCTGTCTTTG 638

Db 61 ggtcaagtcttaacctgcacctgagttgctcctctgcacccacacacctctg 120

Qy 639 GGTCAAGTCTTAATCCCTGCACCTGAGTTGGTCTCTGACCCACACCTCTG 698

Db 121 ccgntntgcaactggaagaggagtgctgatttaagcctttgcgctcgggg 180

Qy 699 CCCGCTGGCACTGGAAGAGGAGTGGCTGATTTAAGCCCTTTGCCCTCGGG 758

Db 181 accagcagaactcctggcagcagctgctcttgtagaagaacttagatacctctc 240

Qy 759 ACCAGCACATCTTGGGACGCCAGCTGCTTTGTAGAGAGACTTAGATACCTCTC 818

Db 241 actctgtttcttgcctccaccccggnatgacagntgtcctctgg 291

Qy 819 ACTTCTGTTCTTCCCTGCCACCCCGGCCATGCCAGTGTGCTCCTCGG 869

RESULT 2 AA295958 279 bp mRNA EST 18-APR-1997

LOCUS EST101182 Thymus III Homo sapiens cDNA 5' end.

DEFINITION AA295958

ACCESSION 91948303

NID EST.

KEYWORDS human.

SOURCE Homo sapiens

ORGANISM Eukaryotes; Metazoa; Chordata;

Vertebrata; Mammalia; Euthera; Primates; Catarrhini; Homiidae;

Homo.

REFERENCE

AUTHORS

Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A., White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man-Wai, C., Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D., Fitzgerald, L.M., Fitzhugh, W.M., Fritchman, J.L., Geohagen, N.S., Glodok, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr., Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M., Moreno-Palancas, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M., Phillips, C.A., Ryder, S.E., Scott, J.L., Saudek, D.M., Shirley, R., Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y., Bednarik, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J., Dimke, D., Feng, D.-F., Ferrie, A., Fischer, C., Hastings, G.A., He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Kozak, D.L., Kunsch, C., Hungjun, J., Li, H., Weissner, P.S., Olsen, H., Raymond, L., Wei, Y.F., Wing, J., Xu, C., Yu, G.L., Ruben, S.M., Dillion, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M. and Venter, J.C.

TITLE Initial assessment-of-human-gene-diversity and expression patterns based upon 83 million nucleotides of cDNA-sequence

JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)

MEDLINE 9602680

COMMENT Other_ESTs: THC148619

Contact: Kerlavage, AR

Bioinformatics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056

Fax: 3018699423

Email: arkerlav@tigr.org

For clone availability, additional sequence and expression

information related to this EST, please check the TIGR Human Gene

Index (<http://www.tigr.org/tdb/hgi/html>)

Seq primer: M13 Reverse.

FEATURES Location/Qualifiers

source

1..279

/organism="Homo sapiens"

/note="Organ: thymus; Vector: pBluescript SK-; Site_1:

ECORI; Site_2: XhoI"

/clone_lib="Thymus III"

/dev_stage="adult"

<1..>279

BASE COUNT 48 a 85 c 62 g 78 t 6 others

ORIGIN

Query Match 26.4%; Score 261; DB 25; Length 279;

Best Local Similarity 95.7%; Pred. No. 0.00e+00;

Matches 267; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Db 1 agcctcttttttggntcaagnttaacccctgcacctaanttgctcctcctctgca 60

Qy 623 AGCGCTCTGTCTTTTGGTCAAGTCTTAATCCTCGACCTGAGTTGGTCTCCTCTGCA 682

Db 61 ccccccacacactcctgcgcttggtgcaactggaagaggaggttgctgattnaagcc 120

Qy 683 CCCCCACCACTCTCTGCCGCTGCACTGGAAAGAGGAGTTGGCTGATTTAAGCC 742

Db 121 ttttccgctcccgagaccagcagcaactcctggcagcagtggtctttagagaagac 180

Qy 743 TTTTGGCGCTCCGGGACCCAGCAGCAATCTGGGACGACGAGTGGCTCTTTAGAGAAGAC 802

Db 181 ttaggatacctctntnaactttctgttttgcgtccacccgggccaagcagttgtc 240

Qy 803 TTAGGATACCTCTCTACACTTTCTGTTTCGCTGTCACCCCGGCCATGCCAGTGTGTC 862

Db 241 cctctgggtccctccaaaactttggtcagttcaagatg 279

Qy 863 CCTCTGGTCCCTCCAAAACCTCTGGTCAGTTCAAGGATG 901

RESULT 3 AA295925 274 bp mRNA EST 18-APR-1997

LOCUS EST101148 Thymus III Homo sapiens cDNA 5' end.

DEFINITION

371 AGGTTTTCGCAAGCT

371 AGGTTTTCGAAAGCTCCACCACAACACGCAGACCTTCCAAG ---GCCCTCTGTGGGTAA 427

ORIGIN

ORIGIN